TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-400-256-19

```
Qy    18 CGACAAAACGGTAAAA 33
         ||||||||||||||||
Db   230 CGACAAAACGGTAAAA 245
```

Query Match   24.2%; Score 16; DB 1; Length 523;
Best Local Similarity   100.0%; Pred. No. 6.4;
Matches   16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-08-400-256-19

RESULT 5
US-08-400-256-34
Sequence 34, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asser Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5750497disk of No. 5750497th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,256
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655

```
Qy    18 CGACAAAACGGTAAAA 33
         ||||||||||||||||
Db   230 CGACAAAACGGTAAAA 245
```

Query Match   24.2%; Score 16; DB 1; Length 523;
Best Local Similarity   100.0%; Pred. No. 6.4;
Matches   16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-08-400-256-34

RESULT 6

US-08-400-256-43
Sequence 43, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asser Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5750497o No. 5750497disk of No. 5750497th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,256
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-400-256-43

```
Qy    18 CGACAAAACGGTAAAA 33
         ||||||||||||||||
Db   230 CGACAAAACGGTAAAA 245
```

Query Match   24.2%; Score 16; DB 1; Length 523;
Best Local Similarity   100.0%; Pred. No. 6.4;
Matches   16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-08-400-256-43

RESULT 7
US-08-975-365-19
Sequence 19, Application US/08975365
Patent No. 6011007
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asser Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

OM nucleic - nucleic search, using sw model

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

Run on: November 5, 2002, 05:43:24 ; Search time 22.2973 Seconds
                727.075 Million cell updates/sec
                (without alignments)

Title: US-09-766-113-2
Perfect score: 66
Sequence: 1 taggatctgaatccgttcga............ccgtcctacgaattcagctg 66

Scoring table: OLIGO NUC
                Gapop 60.0 , Gapext 60.0
                Gapop 60.0 , Gapext 60.0

Searched: 383533 seqs, 122816752 residues

Word size :   0

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   Issued_Patents_NA:*
                1: /cqn2_6/ptodata/1/ina/5A_COMB.seq:*
                2: /cqn2_6/ptodata/1/ina/5B_COMB.seq:*
                3: /cqn2_6/ptodata/1/ina/6A_COMB.seq:*
                4: /cqn2_6/ptodata/1/ina/6B_COMB.seq:*
                5: /cqn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                6: /cqn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	16	24.2	61	4	US-08-563-524A-10	Sequence 10, Appl
2	16	24.2	511	1	US-08-400-256-40	Sequence 40, Appl
3	16	24.2	511	3	US-08-975-365-40	Sequence 40, Appl
4	16	24.2	523	1	US-08-400-256-19	Sequence 19, Appl
5	16	24.2	523	1	US-08-400-256-34	Sequence 34, Appl
6	16	24.2	523	1	US-08-400-256-43	Sequence 43, Appl
7	16	24.2	523	3	US-08-975-365-19	Sequence 19, Appl
8	16	24.2	523	3	US-08-975-365-34	Sequence 34, Appl
9	16	24.2	523	3	US-08-975-365-43	Sequence 43, Appl
10	16	24.2	535	1	US-08-400-256-46	Sequence 46, Appl
11	16	24.2	535	3	US-08-975-365-46	Sequence 46, Appl
12	16	24.2	538	1	US-08-400-256-49	Sequence 49, Appl
13	16	24.2	538	3	US-08-975-365-49	Sequence 49, Appl
14	15	22.7	36	1	US-08-136-993-15	Sequence 15, Appl
15	15	22.7	533	3	US-09-057-570-3	Sequence 3, Appl
16	15	22.7	533	4	US-08-449-083-7	Sequence 7, Appl
17	15	22.7	2964	2	US-08-286-819A-18	Sequence 18, Appl
18	15	22.7	2964	3	US-08-980-357-18	Sequence 18, Appl
19	15	22.7	3190	2	US-08-286-819A-30	Sequence 30, Appl
20	15	22.7	3190	3	US-08-980-357-30	Sequence 30, Appl
21	15	22.7	7004	3	US-09-057-570-3	Sequence 3, Appl
22	15	22.7	10851	2	US-08-286-819A-16	Sequence 16, Appl
23	15	22.7	10851	3	US-08-980-357-16	Sequence 16, Appl
24	14	21.2	24	1	US-08-265-628-10	Sequence 10, Appl
25	14	21.2	39	1	US-08-459-064B-34	Sequence 34, Appl
26	14	21.2	39	2	US-08-460-421A-34	Sequence 34, Appl
27	14	21.2	185	1	US-08-554-369A-1	Sequence 1, Appl
28	14	21.2	1037	4	US-08-858-207A-149	Sequence 149, Appl
29	14	21.2	1353	4	US-09-518-657-3	Sequence 3, Appl
30	14	21.2	1598	4	US-09-518-657-6	Sequence 6, Appl
31	14	21.2	2100	1	US-08-332-576-1	Sequence 1, Appl
32	14	21.2	2100	5	PCT-US95-13672-1	Sequence 1, Appl
33	14	21.2	2359	1	US-08-188-582-4	Sequence 4, Appl
34	14	21.2	2359	4	US-08-644-715-4	Sequence 4, Appl
35	14	21.2	6736	3	US-09-057-570-1	Sequence 1, Appl
36	14	21.2	6736	4	US-09-057-570-5	Sequence 5, Appl
37	14	21.2	19932	2	US-09-150-805-13	Sequence 13, Appl
38	13	19.7	18	3	US-09-150-805-13	Sequence 13, Appl
39	13	19.7	18	4	US-08-996-069A-13	Sequence 13, Appl
40	13	19.7	18	4	US-08-996-069A-13	Sequence 13, Appl
41	13	19.7	18	4	US-09-178-089-13	Sequence 13, Appl
42	13	19.7	22	4	US-09-178-089-13	Sequence 13, Appl
43	13	19.7	24	4	US-08-480-173A-25	Sequence 25, Appl
44	13	19.7	24	3	US-08-480-173A-26	Sequence 26, Appl
45	13	19.7	24	3	US-08-484-408A-25	Sequence 25, Appl

ALIGNMENTS

RESULT 1
US-08-563-524A-10

Sequence 10, Application US/0856 3524A
Patent No. 6232095

GENERAL INFORMATION:
  APPLICANT: Kmiec et al.
  TITLE OF INVENTION: Recombinant Helix Modification
  TITLE OF INVENTION: Recognition Proteins and Uses Thereof
  NUMBER OF SEQUENCES: 32
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Jane Massey Licata, Esq.
  STREET: 210 Lake Drive East, Suite 201
  CITY: Cherry Hill
  STATE: NJ
  COUNTRY: USA
  ZIP: 08002
  COMPUTER READABLE FORM:
  MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
  COMPUTER: IBM 486
  OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
  SOFTWARE: WORDPERFECT 5.1
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/563,524A
  FILING DATE: NO. 6232095ember 28, 1995
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
  APPLICATION NUMBER:
  FILING DATE:
  ATTORNEY/AGENT INFORMATION:
  NAME: Jane Massey Licata
  REGISTRATION NUMBER: 32,257
  REFERENCE/DOCKET NUMBER: TJU-1390
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: (609) 779-2400
  TELEFAX: (609) 779-8488
  INFORMATION FOR SEQ ID NO: 10:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 61
  TYPE: Nucleic Acid
  STRANDEDNESS: Single
  TOPOLOGY: Linear
  ANTI-SENSE: NO
US-08-563-524A-10

Query Match   24.2%; Score 16; DB 4; Length 61;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   50 TCCTACGAATTCAGCT 65
                  ||||||||||||||||

Db   47 TCCTACGAATTCAGTT 52

RESULT 2
US-08-400-256-40
Sequence 40, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5750497disk of No. 5750497th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,256
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212 878-9655
INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 511 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-400-256-40

Query Match   24.2%; Score 16; DB 1; Length 511;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   18 CGACAAAACGGTAAAA 33
         ||||||||||||||||
Db   227 CGACAAAACGGTAAAA 242

RESULT 3
US-08-975-365-40
Sequence 40, Application US/08975365
Patent No. 6011007
GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 511 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-975-365-40

Query Match   24.2%; Score 16; DB 3; Length 511;
Best Local Similarity 100.0%; Pred. No. 6.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   18 CGACAAAACGGTAAAA 33
         ||||||||||||||||
Db   227 CGACAAAACGGTAAAA 242

RESULT 4
US-08-400-256-19
Sequence 19, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5750497disk of No. 5750497th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,256
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-400-256-19

Query Match
Best Local Similarity   24.2%;   Score 16;   DB 1;   Length 523;
Matches   16;   Conservative   0;   Pred. No. 6.4;   Mismatches   0;   Indels   0;   Caps   0;

Qy   18 CGACAAAACGGTAAAA 33
          ||||||||||||||||
Db   230 CGACAAAACGGTAAAA 245

RESULT 5
US-08-400-256-34
Sequence 34, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asser Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5750497disk of No. 5750497th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,256
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-400-256-34

Query Match
Best Local Similarity   24.2%;   Score 16;   DB 1;   Length 523;
Matches   16;   Conservative   0;   Pred. No. 6.4;   Mismatches   0;   Indels   0;   Caps   0;

Qy   18 CGACAAAACGGTAAAA 33
          ||||||||||||||||
Db   230 CGACAAAACGGTAAAA 245

RESULT 6

US-08-400-256-43
Sequence 43, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asser Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5750497disk of No. 5750497th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/400,256
FILING DATE: 03-MAR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Lambiris, Elias J.
REGISTRATION NUMBER: 33,728
REFERENCE/DOCKET NUMBER: 3985.220-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-867-0123
TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 43:
SEQUENCE CHARACTERISTICS:
LENGTH: 523 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-400-256-43

Query Match
Best Local Similarity   24.2%;   Score 16;   DB 1;   Length 523;
Matches   16;   Conservative   0;   Pred. No. 6.4;   Mismatches   0;   Indels   0;   Caps   0;

Qy   18 CGACAAAACGGTAAAA 33
          ||||||||||||||||
Db   230 CGACAAAACGGTAAAA 245

RESULT 7
US-08-975-365-19
Sequence 19, Application US/08975365
Patent No. 6011007
GENERAL INFORMATION:
APPLICANT: Havelund, Svend
APPLICANT: Halstrom, John
APPLICANT: Jonassen, Ib
APPLICANT: Andersen, Asser Sloth
APPLICANT: Markussen, Jan
TITLE OF INVENTION: ACYLATED INSULIN
NUMBER OF SEQUENCES: 49
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
STREET: 405 Lexington Avenue, 64th Floor
CITY: New York
STATE: New York
COUNTRY: United States of America
ZIP: 10174-6401
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985,220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-975-465-19

Query Match   24.2%;   Score 16; DB 3;   Length 523;
Best Local Similarity 100.0%;   Pred. No. 6.4;
Matches 16;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY   18 CGACAAAACGGTAAAA 33
        ||||||||||||||||
DB 230 CGACAAAACGGTAAAA 245

RESULT 8
US-08-975-465-44
; Sequence 44, Application US/08975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985,220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 34:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-975-365-34

Query Match   24.2%;   Score 16; DB 3;   Length 523;
Best Local Similarity 100.0%;   Pred. No. 6.4;
Matches 16;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY   18 CGACAAAACGGTAAAA 33
        ||||||||||||||||
DB 230 CGACAAAACGGTAAAA 245

RESULT 9
US-08-975-465-43
; Sequence 43, Application US/08975365
; Patent No. 6011007
; GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985,220-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-878-9655
; TELEFAX: 212-867-0123
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-08-975-365-43

Query Match   24.2%;   Score 16; DB 3;   Length 523;
Best Local Similarity 100.0%;   Pred. No. 6.4;
Matches 16;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

QY   18 CGACAAAACGGTAAAA 33
        ||||||||||||||||
DB 230 CGACAAAACGGTAAAA 245

RESULT 10
US-08-400-256-46
Sequence 46, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5750497disk of No. 5750497th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,256
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 535 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-400-256-46

Query Match   24.28;   Score 16; DB 1; Length 535;
Best Local Similarity   100.0%;   Pred. No. 6.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   18 CGACAAAACGGTAAAA 33
        ||||||||||||||||
Db 227 CGACAAAACGGTAAAA 242

RESULT 11
US-08-975-365-46
Sequence 46, Application US/08975365
Patent No. 6011007
GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6011007o No. 6011007disk of No. 6011007th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,365
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/400,256
; FILING DATE: 03-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 535 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-975-365-46

Query Match   24.28;   Score 16; DB 3; Length 535;
Best Local Similarity   100.0%;   Pred. No. 6.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   18 CGACAAAACGGTAAAA 33
        ||||||||||||||||
Db 227 CGACAAAACGGTAAAA 242

RESULT 12
US-08-400-256-49
Sequence 49, Application US/08400256
Patent No. 5750497
GENERAL INFORMATION:
; APPLICANT: Havelund, Svend
; APPLICANT: Halstrom, John
; APPLICANT: Jonassen, Ib
; APPLICANT: Andersen, Asser Sloth
; APPLICANT: Markussen, Jan
; TITLE OF INVENTION: ACYLATED INSULIN
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5750497o No. 5750497disk of No. 5750497th America, Inc.
; STREET: 405 Lexington Avenue, 64th Floor
; CITY: New York
; STATE: New York
; COUNTRY: United States of America
; ZIP: 10174-6401
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/400,256
; FILING DATE: 03-MAR-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 3985.220-US
; TELEPHONE: 212-867-0123
; TELEFAX: 212-878-9655
; INFORMATION FOR SEQ ID NO: 49:

Query Match   24.28;   Score 16; DB 3; Length 535;
Best Local Similarity   100.0%;   Pred. No. 6.4;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   18 CGACAAAACGGTAAAA 33
        ||||||||||||||||
Db 227 CGACAAAACGGTAAAA 242

SEQUENCE CHARACTERISTICS:
  LENGTH: 538 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: DNA
US-08-400-256-49

Qy 18 CUACAAAAGGTTAAAA 33
       ||||||||||||||||||
Db 227 CUACAAAAGGTTAAAA 242

Query Match   24.28%; Score 16; DB 1; Length 538;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13
US-08-975-465-49
Sequence 49, Application US/08975465
Patent No. 6011007
GENERAL INFORMATION:
  APPLICANT: Havelund, Svend
  APPLICANT: Halstrom, John
  APPLICANT: Jonassen, Ib
  APPLICANT: Andersen, Asser Sloth
  APPLICANT: Markussen, Jan
  TITLE OF INVENTION: ACYLATED INSULIN
  NUMBER OF SEQUENCES: 49
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: No. 6011007o No. 6011007disk ot No. 6011007th America, Inc.
    STREET: 405 Lexington Avenue, 64th Floor
    CITY: New York
    STATE: New York
    COUNTRY: United States of America
    ZIP: 10174-6401
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/975,365
    FILING DATE:
    CLASSIFICATION: 514
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/400,256
    FILING DATE: 04-MAR-1995
  ATTORNEY/AGENT INFORMATION:
    NAME: Lambiris, Elias J.
    REGISTRATION NUMBER: 33,728
    REFERENCE/DOCKET NUMBER: 3985.220-US
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 212-867-0123
    TELEFAX: 212-878-9655
INFORMATION FOR SEQ ID NO: 49:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 518 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: DNA
US-08-975-465-49

Qy 18 CUACAAAAGGTTAAAA 33
       ||||||||||||||||||
Db 227 CUACAAAAGGTTAAAA 242

Query Match   24.28%; Score 16; DB 3; Length 538;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 14
US-08-136-993-15/c
Sequence 15, Application US/08136993
Patent No. 5420025
GENERAL INFORMATION:
  APPLICANT: Takaui, Hiroshi
  APPLICANT: Aratika, Shino
  APPLICANT: Matsui, Hiroshi
  APPLICANT: Washizu, Kinya
  APPLICANT: Ando, Keiichi
  APPLICANT: Koikeda, Satoshi
  TITLE OF INVENTION: Recombinant transglutaminase
  NUMBER OF SEQUENCES: 22
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Sughrue, Mion, Zinn, Macpeak & Seas
    STREET: 2100 Pennsylvania Avenue
    CITY: N.W.
    STATE: Washington, D.C.
    COUNTRY: U.S.A.
    ZIP: 20037-3202
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.24
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/136,993
    FILING DATE:
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US/07/777,447
    FILING DATE:
    APPLICATION NUMBER: JP 2-282566
    FILING DATE: 19-OCT-1990
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 202-293-7060
    TELEFAX: 202-293-7860
    TELEX: 6491103
INFORMATION FOR SEQ ID NO: 15:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 36 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: Other nucleic acid (synthetic DNA)
US-08-136-993-15

Qy 1 TGAGATCTGGATCCG 15
      |||||||||||||||
Db 16 TGAGATCTGGATCCG 2

Query Match   22.78%; Score 15; DB 1; Length 36;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15
US-08-735-545-7
Sequence 7, Application US/08735545
Patent No. 6025131
GENERAL INFORMATION:
  APPLICANT: Van Dyk, Tina K.
  APPLICANT: LaRossa, Robert Alan
  TITLE OF INVENTION: A Facile Method for
  TITLE OF INVENTION: Identifying Regulated
  TITLE OF INVENTION: Promoters
  NUMBER OF SEQUENCES: 28
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: E. I. DU PONT DE NEMOURS
    ADDRESSEE: AND COMPANY
    STREET: 1007 MARKET STREET
    CITY: WILMINGTON
    STATE: DELAWARE

COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
SOFTWARE: MICROSOFT WORD 2.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/735,545
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA A.
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9989
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
STRAIN: dpd2087 upper
US-08-735-545-7

```
Qy  22  AAAACGGTAAAAAAG  36
        ||||||||||||||
Db  72  AAAACGGTAAAAAAC  86
```

Query Match   22.7%; Score 15; DB 3; Length 533;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 16
US-09-449-083-7
Sequence 7, Application US/09449083
Patent No. 6194159
GENERAL INFORMATION:
APPLICANT: Van Dyk, Tina K.
APPLICANT: LaRossa, Robert Alan
TITLE OF INVENTION: A Facile Method for
TITLE OF INVENTION: Identifying Regulated
TITLE OF INVENTION: Promoters
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS
ADDRESSEE: AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 3.1
SOFTWARE: MICROSOFT WORD 2.0C
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/449,083
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA A.
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9989
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
STRAIN: dpd2087 upper
US-09-449-083-7

```
Qy  22  AAAACGGTAAAAAAG  36
        ||||||||||||||
Db  72  AAAACGGTAAAAAAC  86
```

Query Match   22.7%; Score 15; DB 4; Length 533;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 17
US-08-286-819A-18
Sequence 18, Application US/0828681 9A
Patent No. 5871910
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 587191 0man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2964 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2964
US-08-286-819A-18

Qy   20 ACAAAACGGTAAAAA 34
         |||||||||||||||
Db   966 ACAAAACGGTAAAAA 980

Query Match   22.7%; Score 15; DB 2; Length 2964;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 18
US-08-980-457-18
Sequence 18, Application US/08980357
Patent No. 6013508
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/980,357
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,819
FILING DATE: 05-AUG-1994
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 6013508man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220
TELEX: 248855 OPAT UR
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2964 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2964
US-08-980-357-18

Qy   20 ACAAAACGGTAAAAA 34
         |||||||||||||||
Db   966 ACAAAACGGTAAAAA 980

Query Match   22.7%; Score 15; DB 3; Length 2964;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 19
US-08-286-819A-30
Sequence 30, Application US/0828681 9A
Patent No. 5871910
GENERAL INFORMATION:
APPLICANT: ARTHUR, MICHEL
APPLICANT: DUKTA-MALEN, SYLVIE
APPLICANT: MOLINAS, CATHERINE
APPLICANT: COURVALIN, PATRICE
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
ADDRESSEE: P.C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,819A
FILING DATE: 05-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/174,682
FILING DATE: 28-DEC-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/917,146
FILING DATE: 10-AUG-1992
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9013579
FILING DATE: 31-OCT-1990
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR/91/00855
FILING DATE: 29-OCT-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Oblon, No. 5871910man F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-060-0 PCT

RESULT 20
US-08-980-357-30
Sequence 40, Application US/08980357
Patent No. 6013508
GENERAL INFORMATION:
  APPLICANT: ARTHUR, MICHEL
  APPLICANT: DUKTA-MALEN, SYLVIE
  APPLICANT: MOLINAS, CATHERINE
  APPLICANT: COURVALIN, PATRICE
  TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
  TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
  TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
  TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
  NUMBER OF SEQUENCES: 54
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: OBLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
    ADDRESSEE: P.C.
    STREET: 1755 S. Jefferson Davis Highway, Suite 400
    CITY: Arlington
    STATE: Virginia
    COUNTRY: U.S.A.
    ZIP: 22202
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/980,357
    FILING DATE:
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/286,819
    FILING DATE: 05-AUG-1994
    APPLICATION NUMBER: US 08/174,682
    FILING DATE: 28-DEC-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/917,146
    FILING DATE: 10-AUG-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: PCT/FR/91/00855
    FILING DATE: 29-OCT-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: FR 9013579
    FILING DATE: 31-OCT-1990
  ATTORNEY/AGENT INFORMATION:
    NAME: Oblon, No. 60135088man F.
    REGISTRATION NUMBER: 24,618
    REFERENCE/DOCKET NUMBER: 660-060-0 PCT
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (703) 413-3000
    TELEFAX: (703) 413-2220
    TELEX: 248855 OPAT UR
  INFORMATION FOR SEQ ID NO: 30:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 3190 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
US-08-286-819A-30

Query Match 22.7%; Score 15; DB 2; Length 3190;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ACAAAACGGTAAAAA 34
       |||||||||||||||
Db 1115 ACAAAACGGTAAAAA 1129

RESULT 21
US-09-057-570-3/c
Sequence 3, Application US/09057570
Patent No. 6013266
GENERAL INFORMATION:
  APPLICANT: Seqers, Ruud P.A.M.
  APPLICANT: Frey, Joachim
  TITLE OF INVENTION: Live attenuated Actinobacillus
  TITLE OF INVENTION: pleuropneumoniae
  NUMBER OF SEQUENCES: 7
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Akzo No. 601326601 Patent Department
    STREET: 1300 Piccard Drive, Suite 206
    CITY: Rockville
    STATE: Maryland
    COUNTRY: USA
    ZIP: 20850
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/057,570
    FILING DATE: 09-APR-1998
  ATTORNEY/AGENT INFORMATION:
    NAME: Gormley, Mary E.
    REGISTRATION NUMBER: 34,409
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (301) 948-7400
    TELEFAX: (301) 948-9751
  INFORMATION FOR SEQ ID NO: 3:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 7004 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    HYPOTHETICAL: NO
    ANTI-SENSE: NO
    ORIGINAL SOURCE:
      ORGANISM: Actinobacillus pleuropneumoniae
      STRAIN: HV114 (serotype 3 field strain)
    IMMEDIATE SOURCE:
      CLONE: pROK5
    FEATURE:
      NAME/KEY: CDS
      LOCATION: 1566..5714
      OTHER INFORMATION: /codon_start= 1566
      OTHER INFORMATION: /function= "RTX-toxin"
      OTHER INFORMATION: /product= "ApxIV_var3"
      OTHER INFORMATION: /gene= "apxIV_var3"
      OTHER INFORMATION: /number= 1
    TELEFAX: (703) 413-2220
    TELEX: 248855 OPAT UR
  INFORMATION FOR SEQ ID NO: 30:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 3190 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
US-08-980-357-30

Query Match 22.7%; Score 15; DB 3; Length 3190;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 20 ACAAAACGGTAAAAA 34
       |||||||||||||||
Db 1115 ACAAAACGGTAAAAA 1129

US-09-057-570-4

Query Match   22.7%;   Score 15;   DB 3;   Length 7004;
Best Local Similarity   100.0%;   Pred. No. 15;
Matches 15;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

QY   29 TAAAAAGCGGTAGA 43
              ||||||||||||||
Db   542 TAAAAAGCGGTAGA 528

RESULT 22
US-08-286-819A-16/c:
Sequence 16, Application US/08286819A
Patent No. 5871910
GENERAL INFORMATION:
  APPLICANT: ARTHUR, MICHEL
  APPLICANT: DUKTA-MALEN, SYLVIE
  APPLICANT: MOLINAS, CATHERINE
  APPLICANT: COURVALIN, PATRICE
  TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
  TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN THE
  TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
  TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
  NUMBER OF SEQUENCES: 54
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: ORLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
    ADDRESSEE: P.C.
    STREET: 1755 S. Jefferson Davis Highway, Suite 400
    CITY: Arlington
    STATE: Virginia
    COUNTRY: U.S.A.
    ZIP: 22202
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/286,819A
    FILING DATE: 05-AUG-1994
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/174,682
    FILING DATE: 28-DEC-1993
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/917,146
    FILING DATE: 10-AUG-1992
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: PCT/FR/91/00855
    FILING DATE: 29-OCT-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: FR 9013579
    FILING DATE: 31-OCT-1990
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Oblon, No. 5871910man F.
    REGISTRATION NUMBER: 24,618
    REFERENCE/DOCKET NUMBER: 660-060-0 PCT
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (703) 413-3000
    TELEFAX: (703) 413-2220
    TELEX: 248855 OPAT UR
  INFORMATION FOR SEQ ID NO: 16:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 10851 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: unknown
  MOLECULE TYPE: DNA (genomic)
US-08-286-819A-16

Query Match   22.7%;   Score 15;   DB 2;   Length 10851;
Best Local Similarity   100.0%;   Pred. No. 14;
Matches 15;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

QY   20 ACAAAACGGTAAAAA 34
              |||||||||||||||
Db   2076 ACAAAACGGTAAAAA 2062

RESULT 23
US-08-980-357-16/c:
Sequence 16, Application US/08980357
Patent No. 6013508
GENERAL INFORMATION:
  APPLICANT: ARTHUR, MICHEL
  APPLICANT: DUKTA-MALEN, SYLVIE
  APPLICANT: MOLINAS, CATHERINE
  APPLICANT: COURVALIN, PATRICE
  TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
  TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
  TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
  TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
  NUMBER OF SEQUENCES: 54
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: ORLON, SPIVAK, McCLELLAND, MAIER & NEUSTADT,
    ADDRESSEE: P.C.
    STREET: 1755 S. Jefferson Davis Highway, Suite 400
    CITY: Arlington
    STATE: Virginia
    COUNTRY: U.S.A.
    ZIP: 22202
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/980,357
    FILING DATE:
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/286,819
    FILING DATE: 05-AUG-1994
    APPLICATION NUMBER: US 08/174,682
    FILING DATE: 28-DEC-1993
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 07/917,146
    FILING DATE: 10-AUG-1992
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: PCT/FR/91/00855
    FILING DATE: 29-OCT-1991
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: FR 9013579
    FILING DATE: 31-OCT-1990
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Oblon, No. 6013508man F.
    REGISTRATION NUMBER: 24,618
    REFERENCE/DOCKET NUMBER: 660-060-0 PCT
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (703) 413-3000
    TELEFAX: (703) 413-2220
    TELEX: 248855 OPAT UR
  INFORMATION FOR SEQ ID NO: 16:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 10851 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: unknown
  MOLECULE TYPE: DNA (genomic)
US-08-980-357-16

Query Match   22.7%;   Score 15;   DB 3;   Length 10851;

Host Local Similarity 100.0%; Pred. No. 14; Mismatches 0; Indels 0; Gaps 0;
Matches 15; Conservative 0;

```
Qy    20 ACAAAACGGTAAAAA 34
         |||||||||||||||
Db  2076 ACAAAACGGTAAAAA 2062
```

RESULT 24
US-08-265-628-10/c
Sequence 10, Application US/08265628
Patent No. 5821094
GENERAL INFORMATION:
   APPLICANT: Rothstein, Steven J.
   APPLICANT: Goring, Daphne
   TITLE OF INVENTION: S-LOCUS RECEPTOR KINASE GENE IN A
   TITLE OF INVENTION: SELF-INCOMPATIBLE BRASSICA NAPUS LINE
   NUMBER OF SEQUENCES: 10
   CORRESPONDENCE ADDRESS:
   ADDRESSEE: McAndrews, Held & Malloy, Ltd.
   STREET: 500 W. Madison St., Suite 3400
   CITY: Chicago
   STATE: Illinois
   COUNTRY: USA
   ZIP: 60661
   COMPUTER READABLE FORM:
   MEDIUM TYPE: Floppy disk
   COMPUTER: IBM PC compatible
   OPERATING SYSTEM: PC-DOS/MS-DOS
   SOFTWARE: PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
   APPLICATION NUMBER: US/08/265,628
   FILING DATE:
   PRIOR APPLICATION DATA:
   APPLICATION NUMBER: US/07/959,945
   FILING DATE:
   APPLICATION NUMBER: US 07/847,564
   FILING DATE: 03-MAR-1992
   ATTORNEY/AGENT INFORMATION:
   NAME: Pochopien Ph.D., Donald J.
   REGISTRATION NUMBER: 32,167
   TELECOMMUNICATION INFORMATION:
   TELEPHONE: 312-707-8889
   TELEFAX: 312-707-9155
   INFORMATION FOR SEQ ID NO: 10:
   SEQUENCE CHARACTERISTICS:
   LENGTH: 24 base pairs
   TYPE: nucleic acid
   STRANDEDNESS: single
   TOPOLOGY: linear
   MOLECULE TYPE: cDNA
   HYPOTHETICAL: NO
   ANTI-SENSE: NO
   FEATURE:
   NAME/KEY: misc_feature
   LOCATION: 1..24

US-08-265-628-10

Query Match 21.2%; Score 14; DB 1; Length 24;
Host Local Similarity 100.0%; Pred. No. 1.1e-02; Mismatches 0; Indels 0; Gaps 0;
Matches 14; Conservative 0;

```
Qy   2 GAGATCTGGATCCG 15
       ||||||||||||||
Db  15 GAGATCTGGATCCG 2
```

RESULT 25
US-08-459-064B-34/c
Sequence 34, Application US/08459064B
Patent No. 5747452
GENERAL INFORMATION:

   APPLICANT: RUOSLAHTI, ERKKI I.
   APPLICANT: MORLA, ALEX
   TITLE OF INVENTION: A METHOD OF MODULATING TUMOR CELL MIGRATION
   TITLE OF INVENTION: USING FIBRONECTIN TYPE III PEPTIDES
   NUMBER OF SEQUENCES: 35
   CORRESPONDENCE ADDRESS:
   ADDRESSEE: CAMPBELL & FLORES LLP
   STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
   CITY: SAN DIEGO
   STATE: CALIFORNIA
   COUNTRY: UNITED STATES
   ZIP: 92122
   COMPUTER READABLE FORM:
   MEDIUM TYPE: Floppy disk
   COMPUTER: IBM PC compatible
   OPERATING SYSTEM: PC-DOS/MS-DOS
   SOFTWARE: PatentIn Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
   APPLICATION NUMBER: US/08/459,064B
   FILING DATE: 01-JUN-1995
   CLASSIFICATION: 514
   PRIOR APPLICATION DATA:
   APPLICATION NUMBER: US 08/021,626
   FILING DATE: 16-FEB-1993
   PRIOR APPLICATION DATA:
   APPLICATION NUMBER: US 07/829,462
   FILING DATE: 31-JAN-1992
   PRIOR APPLICATION DATA:
   APPLICATION NUMBER: US 08/340,812
   FILING DATE: 17-NOV-1994
   ATTORNEY/AGENT INFORMATION:
   NAME: CAMPBELL, CATHRYN A.
   REGISTRATION NUMBER: 31,815
   REFERENCE/DOCKET NUMBER: P-LA 1543
   TELECOMMUNICATION INFORMATION:
   TELEPHONE: 619-535-9001
   TELEFAX: 619-535-8949
   INFORMATION FOR SEQ ID NO: 34:
   SEQUENCE CHARACTERISTICS:
   LENGTH: 39 base pairs
   TYPE: nucleic acid
   STRANDEDNESS: single
   TOPOLOGY: circular
   FEATURE:
   NAME/KEY: CDS
   LOCATION: 1..37

US-08-459-064B-34

Query Match 21.2%; Score 14; DB 1; Length 39;
Host Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 14; Conservative 0;

```
Qy   1 TGAGATCTGGATCC 14
       ||||||||||||||
Db  20 TGAGATCTGGATCC 7
```

RESULT 26
US-08-460-421A-34/c
Sequence 34, Application US/0846042lA
Patent No. 5837813
GENERAL INFORMATION:
   APPLICANT: RUOSLAHTI, ERKKI I.
   APPLICANT: MORLA, ALEX
   TITLE OF INVENTION: FIBRONECTIN BINDING SITES AND METHODS OF
   TITLE OF INVENTION: MODULATING FIBRONECTIN EXTRACELLULAR MATRIX ASSEMBLY
   NUMBER OF SEQUENCES: 35
   CORRESPONDENCE ADDRESS:
   ADDRESSEE: CAMPBELL & FLORES LLP
   STREET: 4370 LA JOLLA VILLAGE DRIVE, STE 700
   CITY: SAN DIEGO
   STATE: CALIFORNIA
   COUNTRY: UNITED STATES

ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/460,421A
FILING DATE: 01-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/829,462
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/021,626
FILING DATE: 16-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/440,812
FILING DATE: 17-NOV-1994
ATTORNEY/AGENT INFORMATION:
NAME: CAMPBELL, CATHRYN A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-1A 1542
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-535-9001
TELEFAX: 619-535-8949

INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 39 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
FEATURE:
NAME/KEY: CDS
LOCATION: 1..47

US-08-460-421A-34

Qy   1 TCAGATCTGATCC 14
        |||||||||||||
Db 20 TCAGATCTGATCC 7

Query Match   21.2%;   Score 14;   DB 2;   Length 39;
Best Local Similarity 100.0%;   Pred. No. 1e+02;
Matches 14;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

RESULT 27
US-08-554-469A-1
Sequence 1, Application US/08554469A
Patent No. 5824519
GENERAL INFORMATION:
APPLICANT: NORRIS, JAMES S.
APPLICANT: CLAWSON, GARY A.
TITLE OF INVENTION: TISSUE SPECIFIC AND TARGET RNA SPECIFIC
TITLE OF INVENTION: RIBOZYMES
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESS:
ADDRESSEE: NEEDLE & ROSENBERG, P.C.
STREET: 127 Peachtree Street, Suite 1200
CITY: Atlanta
STATE: Georgia
COUNTRY: USA
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/554,469A
FILING DATE:
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:

NAME: Spratt, Gwendolyn D.
REGISTRATION NUMBER: 36,016
REFERENCE/DOCKET NUMBER: 19070.0030
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404/688-0770
TELEFAX: 404/688-9880
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: oligonucleotide

US-08-554-469A-1

Qy   3 AGATCTGGATCCGT 16
         ||||||||||||||
Db 112 AGATCTGGATCCGT 125

Query Match   21.2%;   Score 14;   DB 1;   Length 185;
Best Local Similarity 100.0%;   Pred. No. 80;
Matches 14;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

RESULT 28
US-08-858-207A-149
Sequence 149, Application US/08858207A
Patent No. 6348328
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Nicholas, Richard
APPLICANT: Stodola, Robert
TITLE OF INVENTION: No. 6348328e1 compounds
NUMBER OF SEQUENCES: 552
CORRESPONDENCE ADDRESS:
ADDRESSEE: SmithKline Beecham Corporation
STREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/858,207A
FILING DATE: 09-MAY-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/017670
FILING DATE: 14-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimmi, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50475
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:

INFORMATION FOR SEQ ID NO: 149:
SEQUENCE CHARACTERISTICS:
LENGTH: 1037 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

US-08-858-207A-149

Query Match   21.2%;   Score 14;   DB 4;   Length 1047;
Best Local Similarity 100.0%;   Pred. No. 61;

Matches   14; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     51 CCTACGAATTCAGC 64
          ||||||||||||||
Db    595 CCTACGAATTCAGC 608
```

RESULT 29
US-09-518-657-3
Sequence 3, Application US/09518657
Patent No. 6335188
GENERAL INFORMATION:
APPLICANT: Schardl, Christopher L.
APPLICANT: Wang, Jinghong
TITLE OF INVENTION: Endophyte Ergot Alkaloid Synthetic Compounds, Compounds
TITLE OF INVENTION: Which Encode Therefor and Related Methods
FILE REFERENCE: P-1060
CURRENT APPLICATION NUMBER: US/09/518,657
CURRENT FILING DATE: 2000-03-03
EARLIER APPLICATION NUMBER: 60/125,490
EARLIER FILING DATE: 1999-03-22
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
  LENGTH: 1353
  TYPE: DNA
  ORGANISM: Neotyphodium coenophialum
US-09-518-657-3

Query Match   21.2%; Score 14; DB 4; Length 1353;
Best Local Similarity 100.0%; Pred. No. 61;
Matches   14; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      5 AICTGGATCCGTTC 18
          ||||||||||||||
Db    338 AICTGGATCCGTTC 351
```

RESULT 30
US-09-518-657-6
Sequence 6, Application US/09518657
Patent No. 6335188
GENERAL INFORMATION:
APPLICANT: Schardl, Christopher L.
APPLICANT: Wang, Jinghong
TITLE OF INVENTION: Endophyte Ergot Alkaloid Synthetic Compounds, Compounds
TITLE OF INVENTION: Which Encode Therefor and Related Methods
FILE REFERENCE: P-1060
CURRENT APPLICATION NUMBER: US/09/518,657
CURRENT FILING DATE: 2000-03-03
EARLIER APPLICATION NUMBER: 60/125,490
EARLIER FILING DATE: 1999-03-22
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
  LENGTH: 1598
  TYPE: DNA
  ORGANISM: Neotyphodium coenophialum
US-09-518-657-6

Query Match   21.2%; Score 14; DB 4; Length 1598;
Best Local Similarity 100.0%; Pred. No. 59;
Matches   14; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy      5 AICTGGATCCGTTC 18
          ||||||||||||||
Db    362 AICTGGATCCGTTC 375
```

RESULT 31
US-08-332-576-1
Sequence 1, Application US/08332576
Patent No. 5756105

GENERAL INFORMATION:
APPLICANT: Weiser, Jeffrey M.
TITLE OF INVENTION: Vaccines for Haemophilis Influenza
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5756105ris
  STREET: One Liberty Place - 46th Floor
  CITY: Philadelphia
  STATE: PA
  COUNTRY: U.S.A.
  ZIP: 19103
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/332,576
  FILING DATE:
  CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
  NAME: Ralph, Rebecca L.
  REGISTRATION NUMBER: 35,152
  REFERENCE/DOCKET NUMBER: CH-536
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 215-568-3100
  TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 2100 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: unknown
  MOLECULE TYPE: DNA (genomic)
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 1553..2005
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 209..1492
US-08-332-576-1

Query Match   21.2%; Score 14; DB 1; Length 2100;
Best Local Similarity 100.0%; Pred. No. 57;
Matches   14; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy     30 AAAAAAGGGTAGA 43
          |||||||||||||
Db   1025 AAAAAAGGGTAGA 1038
```

RESULT 32
PCT-US95-13672-1
Sequence 1, Application PC/TUS9513672
GENERAL INFORMATION:
APPLICANT: Weiser, Jeffrey M.
TITLE OF INVENTION: Vaccines for Haemophilis Influenza
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and
  ADDRESSEE: Norris
  STREET: One Liberty Place - 46th Floor
  CITY: Philadelphia
  STATE: PA
  COUNTRY: U.S.A.
  ZIP: 19103
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: PCT/US95/13672

FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: 08/342,576
  FILING DATE: October 31, 1994
ATTORNEY/AGENT INFORMATION:
  NAME: Ralph, Rebecca L.
  REGISTRATION NUMBER: 35,152
  REFERENCE/DOCKET NUMBER: CH-536
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 215-568-3100
  TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 1:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 2100 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: unknown
  MOLECULE TYPE: DNA (genomic)
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 1554..2005
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 209..1492

PCT-US95-14672-1

```
Qy     30  AAAAAAGCGGTAGA  44
               ||||||||||||||
Db   1025  AAAAAAGCGGTAGA  1038
```

Query Match 21.2%; Score 14; DB 5; Length 2100;
Host Local Similarity 100.0%; Pred. No. 57;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT #4
US-08-188-582-4/c
Sequence 4, Application US/08188582
Patent No. 5534410
GENERAL INFORMATION:
  APPLICANT: Tjian, Robert
  APPLICANT: Comai, Lucio
  APPLICANT: Dynlact, Brian D.
  APPLICANT: Hoey, Timothy
  APPLICANT: Ruppert, Siegfried
  APPLICANT: Tanese, Naoko
  APPLICANT: Wang, Edith
  APPLICANT: Weinzierl, Robert O.J.
  TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
  TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
  NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
  ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
  STREET: 4 Embarcadero Center, Suite 3400
  CITY: San Francisco
  STATE: California
  COUNTRY: USA
  ZIP: 94111-4187
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/188,582
  FILING DATE: 28-JAN-1994
  CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
  NAME: Osman, Richard A
  REGISTRATION NUMBER: 36,627
  REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 4:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 2359 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear
  MOLECULE TYPE: cDNA
  FEATURE:
    NAME/KEY: CDS
    LOCATION: 49..2160

US-08-582-4

```
Qy      9  GGATCCGTTCGACA  22
               ||||||||||||||
Db   1875  GGATCCGTTCGACA  1862
```

Query Match 21.2%; Score 14; DB 1; Length 2359;
Host Local Similarity 100.0%; Pred. No. 56;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT #4
US-08-646-715-4/c
Sequence 4, Application US/08646715
Patent No. 5637686
GENERAL INFORMATION:
  APPLICANT: Tjian, Robert
  APPLICANT: Comai, Lucio
  APPLICANT: Dynlact, Brian D.
  APPLICANT: Hoey, Timothy
  APPLICANT: Ruppert, Siegfried
  APPLICANT: Tanese, Naoko
  APPLICANT: Wang, Edith
  APPLICANT: Weinzierl, Robert O.J.
  TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
  TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
  NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
  ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
  STREET: 4 Embarcadero Center, Suite 3400
  CITY: San Francisco
  STATE: California
  COUNTRY: USA
  ZIP: 94111-4187
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/646,715
  FILING DATE: 09-MAY-1996
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/188,582
  FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
  NAME: Osman, Richard A
  REGISTRATION NUMBER: 36,627
  REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (415) 781-1989
  TELEFAX: (415) 398-3249
  TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 4:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 2359 base pairs
    TYPE: nucleic acid
    STRANDEDNESS: double
    TOPOLOGY: linear

MOLECULE TYPE: cDNA
FEATURE:
  NAME/KEY: CDS
  LOCATION: 49..2160
US-08-646-715-4

Qy   9 GGATCCGTTCGACA 22
        |||||||||||||
Db 1875 GGATCCGTTCGACA 1862

Query Match   21.2%; Score 14; DB 1; Length 2359;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 35
US-09-057-570-1/c
Sequence 1, Application US/09057570
Patent No. 6013266
GENERAL INFORMATION:
  APPLICANT: Segers, Ruud P.A.M.
  APPLICANT: Frey, Joachim
  TITLE OF INVENTION: Live attenuated Actinobacillus
  TITLE OF INVENTION: pleuropneumoniae
  NUMBER OF SEQUENCES: 7
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Akzo No. 6013266e1 Patent Department
    STREET: 1300 Piccard Drive, Suite 206
    CITY: Rockville
    STATE: Maryland
    COUNTRY: USA
    ZIP: 20850
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/057,570
    FILING DATE: 09-APR-1998
  ATTORNEY/AGENT INFORMATION:
    NAME: Gormley, Mary E.
    REGISTRATION NUMBER: 34,409
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (301) 948-7400
    TELEFAX: (301) 948-9751
  INFORMATION FOR SEQ ID NO: 1:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 6736 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: circular
    MOLECULE TYPE: DNA (genomic)
    ORIGINAL SOURCE:
      ORGANISM: Actinobacillus pleuropneumoniae
      STRAIN: 4074 (serotype 1 reference strain)
    IMMEDIATE SOURCE:
      CLONE: pROK7
    FEATURE:
      NAME/KEY: CDS
      LOCATION: 1576..6549
      OTHER INFORMATION: /codon_start= 1576
      OTHER INFORMATION: /function= "RTX-toxin"
      OTHER INFORMATION: /product= "ApxIV_var1"
      OTHER INFORMATION: /gene= "apxIV_var1"
      OTHER INFORMATION: /number= 1
    FEATURE:
      NAME/KEY: CDS
      LOCATION: 1576..6549
US-09-057-570-1

Query Match   21.2%; Score 14; DB 3; Length 6736;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   29 TAAAAAGCGGTAG 42

Db 531 TAAAAAGCGGTAG 518

RESULT 36
US-09-057-570-5/c
Sequence 5, Application US/09057570
Patent No. 6013266
GENERAL INFORMATION:
  APPLICANT: Segers, Ruud P.A.M.
  APPLICANT: Frey, Joachim
  TITLE OF INVENTION: Live attenuated Actinobacillus
  TITLE OF INVENTION: pleuropneumoniae
  NUMBER OF SEQUENCES: 7
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Akzo No. 6013266e1 Patent Department
    STREET: 1300 Piccard Drive, Suite 206
    CITY: Rockville
    STATE: Maryland
    COUNTRY: USA
    ZIP: 20850
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/09/057,570
    FILING DATE: 09-APR-1998
  ATTORNEY/AGENT INFORMATION:
    NAME: Gormley, Mary E.
    REGISTRATION NUMBER: 34,409
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (301) 948-7400
    TELEFAX: (301) 948-9751
  INFORMATION FOR SEQ ID NO: 5:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 6736 base pairs
      TYPE: nucleic acid
      STRANDEDNESS: double
      TOPOLOGY: linear
    MOLECULE TYPE: DNA (genomic)
    HYPOTHETICAL: NO
    ANTI-SENSE: NO
    ORIGINAL SOURCE:
      ORGANISM: Actinobacillus pleuropneumoniae
      STRAIN: 4074 (serotype 1 reference strain)
    IMMEDIATE SOURCE:
      CLONE: pROK7
    FEATURE:
      NAME/KEY: CDS
      LOCATION: 1132..6549
      IDENTIFICATION METHOD: experimental
      OTHER INFORMATION: /codon_start= 1132
      OTHER INFORMATION: /function= "RTX toxin"
      OTHER INFORMATION: /product= "ApxIV"
      OTHER INFORMATION: /evidence= EXPERIMENTAL
      OTHER INFORMATION: /gene= "ApxIV_v1"
    FEATURE:
      NAME/KEY: CDS
      LOCATION: 1..453
      OTHER INFORMATION: /codon_start= 1
      OTHER INFORMATION: /function= "Met-G"
      OTHER INFORMATION: /product= "mrp"
      OTHER INFORMATION: /gene= "mrp"
      OTHER INFORMATION: /standard_name= "mrp"
      OTHER INFORMATION: /label= mrp
    FEATURE:
      NAME/KEY: CDS
      LOCATION: -10_signal
      LOCATION: 617..623
      OTHER INFORMATION: /partial
      OTHER INFORMATION: /standard_name= "-10"
      OTHER INFORMATION: /label= -10_s
    FEATURE:
      NAME/KEY: -35_signal

```
;     LOCATION: 594..599
;     OTHER INFORMATION:
;     OTHER INFORMATION: /standard_name=
;  FEATURE:
;     NAME/KEY: promoter
;     LOCATION: 454..1131
;     OTHER INFORMATION: /function= "promoter"
;     OTHER INFORMATION: /standard_name= "promoter Apx1V"
;     OTHER INFORMATION: /label= promoter
US-09-057-570-5
```

```
Qy   29  TAAAAAAGCGGTAG 42
         |||||||||||||
Db  531  TAAAAAAGCGGTAG 518
```

Query Match 21.2%; Score 14; DB 3; Length 6736;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-08-477-451-25

```
; RESULT 47
; US 08-477-451-25
; Sequence 25, Application US/08477451
; Patent No. 5928865
;  GENERAL INFORMATION:
;     APPLICANT: Covacci, Antonello
;     TITLE OF INVENTION: Helicobacter pylori CagI Region
;     NUMBER OF SEQUENCES: 46
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: Chiron Corporation
;        STREET: 4560 Horton Street
;        CITY: Emeryville
;        STATE: CA
;        COUNTRY: USA
;        ZIP: 94608-2916
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: PatentIn Release #1.0, Version #1.40
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/08/477,451
;        FILING DATE: 07-JUN-1995
;        CLASSIFICATION: 435
;     ATTORNEY/AGENT INFORMATION:
;        NAME: McClung, Barbara G.
;        REGISTRATION NUMBER: 33,114
;        REFERENCE/DOCKET NUMBER: 0335.002
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE: 510-601-2708
;        TELEFAX: 510-655-3542
; INFORMATION FOR SEQ ID NO: 25:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 19932 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
US-08-477-451-25
```

```
Qy   29  TAAAAAAGCGGTAG 42
         |||||||||||||
Db  829  TAAAAAAGCGGTAG 842
```

Query Match 21.2%; Score 14; DB 2; Length 19932;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-150-805-4

```
; RESULT 48
; US-09-150-805-4
; Sequence 3, Application US/09150805
; Patent No. 6140080
```

```
;  GENERAL INFORMATION:
;     APPLICANT: Bruce, Wesley
;     APPLICANT: Lu, Guihua
;     TITLE OF INVENTION: PROMOTER ELEMENTS CONFERRING
;     TITLE OF INVENTION: ROOT-PREFERRED GENE EXPRESSION
;     NUMBER OF SEQUENCES: 19
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
;        STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
;        STREET: Box 1000
;        CITY: Johnston
;        STATE: Iowa
;        COUNTRY: USA
;        ZIP: 50131
;     COMPUTER READABLE FORM:
;        MEDIUM TYPE: Floppy disk
;        COMPUTER: IBM PC compatible
;        OPERATING SYSTEM: PC-DOS/MS-DOS
;        SOFTWARE: PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;        APPLICATION NUMBER: US/09/150,805
;        FILING DATE:
;        CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;        APPLICATION NUMBER: US/08/996,069
;        FILING DATE: 22-DEC-1997
;        APPLICATION NUMBER: US 08/649,172
;        FILING DATE: 17-MAY-1996
;     ATTORNEY/AGENT INFORMATION:
;        NAME: Yates, Michael E.
;        REGISTRATION NUMBER: 36,063
;        REFERENCE/DOCKET NUMBER: 0465R
;     TELECOMMUNICATION INFORMATION:
;        TELEPHONE: (515) 248-4800
;        TELEFAX: (515) 248-4844
; INFORMATION FOR SEQ ID NO: 3:
;     SEQUENCE CHARACTERISTICS:
;        LENGTH: 18 base pairs
;        TYPE: nucleic acid
;        STRANDEDNESS: single
;        TOPOLOGY: linear
;     MOLECULE TYPE: other nucleic acid
;        DESCRIPTION: /desc = "oligonucleotide"
;     FEATURE:
;        NAME/KEY: misc_feature
;        LOCATION: 1..18
;        OTHER INFORMATION: /product= "N7914"
US-09-150-805-3
```

```
Qy  5  ATCTGGATCCGTT 17
       |||||||||||||
Db  5  ATCTGGATCCGTT 17
```

Query Match 19.7%; Score 13; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

US-09-150-805-3

```
; RESULT 49
; US-09-150-805-13
; Sequence 13, Application US/09150805
; Patent No. 6140080
;  GENERAL INFORMATION:
;     APPLICANT: Bruce, Wesley
;     APPLICANT: Lu, Guihua
;     TITLE OF INVENTION: PROMOTER ELEMENTS CONFERRING
;     TITLE OF INVENTION: ROOT-PREFERRED GENE EXPRESSION
;     NUMBER OF SEQUENCES: 19
;     CORRESPONDENCE ADDRESS:
;        ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
;        STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
;        STREET: Box 1000
;        CITY: Johnston
```

STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/150,805
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/996,069
FILING DATE: 22-DEC-1997
APPLICATION NUMBER: US 08/649,172
FILING DATE: 17-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Yates, Michael E.
REGISTRATION NUMBER: 36,063
REFERENCE/DOCKET NUMBER: 0465R
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-09-150-805-13

Qy 5 ATCTGGATCCGTT 17
        |||||||||||||
Db 5 ATCTGGATCCGTT 17

Query Match   19.7%; Score 13; DB 3; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 40
US-08-996-069A-3
Sequence 3, Application US/0899606 9A
Patent No. 6228645
GENERAL INFORMATION:
APPLICANT: Lu, Guihua
APPLICANT: Bruce, Wesley
TITLE OF INVENTION: PROMOTER ELEMENTS CONFERRING
TITLE OF INVENTION: ROOT-PREFERRED GENE EXPRESSION
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,069A
FILING DATE: 22-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/649,172
FILING DATE: 17-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Yates, Michael E.
REGISTRATION NUMBER: 36,063
REFERENCE/DOCKET NUMBER: 0465R
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..18
DESCRIPTION: /desc = "oligonucleotide"
OTHER INFORMATION: /product= "N79l3"

Qy 5 ATCTGGATCCGTT 17
        |||||||||||||
Db 5 ATCTGGATCCGTT 17

US-08-996-069A-3

Query Match   19.7%; Score 13; DB 4; Length 18;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 41
US-08-996-069A-13
Sequence 13, Application US/0899606 9A
Patent No. 6228645
GENERAL INFORMATION:
APPLICANT: Lu, Guihua
APPLICANT: Bruce, Wesley
TITLE OF INVENTION: PROMOTER ELEMENTS CONFERRING
TITLE OF INVENTION: ROOT-PREFERRED GENE EXPRESSION
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
CORRESPONDENCE ADDRESS:
ADDRESSEE: PIONEER HI-BRED INTERNATIONAL, INC.
STREET: Darwin Building, 7100 N.W. 62nd Ave., P.O.
STREET: Box 1000
CITY: Johnston
STATE: Iowa
COUNTRY: USA
ZIP: 50131
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
NUMBER OF SEQUENCES: 19
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/996,069A
FILING DATE: 22-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/649,172
FILING DATE: 17-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Yates, Michael E.
REGISTRATION NUMBER: 36,063
REFERENCE/DOCKET NUMBER: 0465R
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 248-4800
TELEFAX: (515) 248-4844
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 18 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

US-08-996-069A-14

Query Match   19.7%;   Score 13;   DB 4;   Length 18;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY   5  ATCTGGATCGTT 17
        |||||||||||||
Db   5  ATCTGGATCGTT 17
```

RESULT 42
US 09 178-089-13/c
Sequence 13, Application US/09178089
Patent No. 6077942
GENERAL INFORMATION:
  APPLICANT: YADAV, NARENDRA S.
  TITLE OF INVENTION: BINARY VIRAL EXPRESSION SYSTEM IN PLANTS
  FILE REFERENCE: CL-1127
  CURRENT APPLICATION NUMBER: US/09/178,089
  CURRENT FILING DATE: 1998-10-23
  EARLIER APPLICATION NUMBER: 60/063,504
  EARLIER FILING DATE: OCTOBER 24, 1997
  NUMBER OF SEQ ID NOS: 14
  SOFTWARE: Microsoft Word Version 7.0A
  SEQ ID NO 13
  LENGTH: 22
  TYPE: DNA
  ORGANISM: Unknown
  FEATURE:
  OTHER INFORMATION: Description of Unknown organism:primers

US-09-178-089-13

Query Match   19.7%;   Score 13;   DB 3;   Length 22;
Host Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY   4  GATCTGGATCCGT 16
        |||||||||||||
Db  14  GATCTGGATCCGT 2
```

RESULT 43
US 08 480-173A-25
Sequence 25, Application US/08480173A
Patent No. 6072049
GENERAL INFORMATION:
  APPLICANT: Thoma, Hans A
  TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE
  NUMBER OF SEQUENCES: 56
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Popovich & Wiles, P.A.
  STREET: 80 S. 8th Street, Suite 1902
  CITY: Minneapolis
  STATE: MN
  COUNTRY: USA
  ZIP: 55402
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patent In Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/480,173A
  FILING DATE: 07-JUN-1995
  CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
  NAME: Popovich, Thomas E
  REGISTRATION NUMBER: 40,099
  REFERENCE/DOCKET NUMBER: MED100*USD4
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: 612-344-8991
  TELEFAX: 612-344-8994

INFORMATION FOR SEQ ID NO: 25:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 24 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: linear
  MOLECULE TYPE: DNA (synthetic)
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 1..4
  OTHER INFORMATION: /note= "Nucleotides 1-4 form a
  OTHER INFORMATION: single-stranded "sticky end""
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 5..24
  OTHER INFORMATION: /note= "Adapter sequence results
  OTHER INFORMATION: from oligonucleotide duplex formation with nucleotides 5-2
  OTHER INFORMATION: SEQ ID NO: 26"

US-08-480-173A-25

Query Match   19.7%;   Score 13;   DB 4;   Length 24;
Host Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY   3  AGATCTGGATCCG 15
        |||||||||||||
Db   6  AGATCTGGATCCG 18
```

RESULT 44
US-08-480-173A-26/c
Sequence 26, Application US/08480173A
Patent No. 6072049
GENERAL INFORMATION:
  APPLICANT: Thoma, Hans A
  TITLE OF INVENTION: HEPATITIS B SURFACE ANTIGEN VACCINE
  NUMBER OF SEQUENCES: 56
  CORRESPONDENCE ADDRESS:
  ADDRESSEE: Popovich & Wiles, P.A.
  STREET: 80 S. 8th Street, Suite 1902
  CITY: Minneapolis
  STATE: MN
  COUNTRY: USA
  ZIP: 55402
  COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: Patent In Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/480,173A
  FILING DATE: 07-JUN-1995
  CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
  NAME: Popovich, Thomas E
  REGISTRATION NUMBER: 40,099
  REFERENCE/DOCKET NUMBER: MED100*USD4
  TELECOMMUNICATION INFORMATION:
  TELEPHONE: 612-344-8991
  TELEFAX: 612-344-8994
INFORMATION FOR SEQ ID NO: 26:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 24 base pairs
  TYPE: nucleic acid
  STRANDEDNESS: double
  TOPOLOGY: linear
  MOLECULE TYPE: DNA (synthetic)
  FEATURE:
  NAME/KEY: misc_feature
  LOCATION: 1..4
  OTHER INFORMATION: /note= "Nucleotides 1-4 form a
  OTHER INFORMATION: single-stranded "sticky end""
  FEATURE:

```
;  NAME/KEY:  misc_feature
;    LOCATION:  5..24
;    OTHER INFORMATION:
;    OTHER INFORMATION:  /note= "Adapter sequence results
;    OTHER INFORMATION:  from oligonucleotide duplex formation with nucleotides 5-24
US-08-480-173A-26                                          SEQ ID NO: 25"
```

```
Qy    3 AGATCTGGATCCG 15
        |||||||||||||
Db   23 AGATCTGGATCCG 11
```

```
Query Match
Best Local Similarity  100.0%;  Pred. No. 3.5e+02;
Matches  13;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
                       19.78;   Score 13;  DB 3;  Length 24;
```

```
RESULT 45
US-08-484-408A-25
Sequence 25, Application US/08484408A
Patent No. 6117653
GENERAL INFORMATION:
  APPLICANT:  Thoma, Hans A
  TITLE OF INVENTION:  HEPATITIS B SURFACE ANTIGEN VACCINE
  NUMBER OF SEQUENCES:  56
  CORRESPONDENCE ADDRESS:
    ADDRESSEE:  Popovich & Wiles, P.A.
    STREET:  80 S. 8th Street, Suite 1902
    CITY:  Minneapolis
    STATE:  MN
    COUNTRY:  USA
    ZIP:  55402
  COMPUTER READABLE FORM:
    MEDIUM TYPE:  Floppy disk
    COMPUTER:  IBM PC compatible
    OPERATING SYSTEM:  PC-DOS/MS-DOS
    SOFTWARE:  PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER:  US/08/484,408A
    FILING DATE:  07-JUN-1995
    CLASSIFICATION:  435
  ATTORNEY/AGENT INFORMATION:
    NAME:  Popovich, Thomas E
    REGISTRATION NUMBER:  30,099
    REFERENCE/DOCKET NUMBER:  MED11003USD4
  TELECOMMUNICATION INFORMATION:
    TELEPHONE:  612-334-8991
    TELEFAX:  612-334-8994
INFORMATION FOR SEQ ID NO:  25:
  SEQUENCE CHARACTERISTICS:
    LENGTH:  24 base pairs
    TYPE:  nucleic acid
    STRANDEDNESS:  double
    TOPOLOGY:  linear
  MOLECULE TYPE:  DNA (synthetic)
  FEATURE:
    NAME/KEY:  misc_feature
    LOCATION:  1..4
    OTHER INFORMATION:  /note= "Nucleotides 1-4 form a
    OTHER INFORMATION:  single-stranded "sticky end""
  FEATURE:
    NAME/KEY:  misc_feature
    LOCATION:  5..24
    OTHER INFORMATION:
    OTHER INFORMATION:  /note= "Adapter sequence results
    OTHER INFORMATION:  from oligonucleotide duplex formation with nucleotides 5-24
US-08-484-408A-25                                          SEQ ID NO: 26"
```

```
Qy    3 AGATCTGGATCCG 15
        |||||||||||||
US-08-484-408A-25
Query Match
Best Local Similarity  100.0%;  Pred. No. 3.5e+02;
Matches  13;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
                       19.78;   Score 13;  DB 3;  Length 24;
```

```
Db    6 AGATCTGGATCCG 18

Search completed: November  5, 2002, 08:26:11
Job time : 37.2973 secs
```

```
REFERENCE   1 (bases 1 to 3032)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 3924 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1..3032
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT      939 a    600 c    422 g   1071 t
ORIGIN

Query Match             24.28;  Score 16;  DB 6;  Length 3032;
Best local Similarity   100.0%;  Pred. No. 65;
Matches   16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy        29 TAAAAAACGCGGTAGAT 44
             |||||||||||||||||
Db      3026 TAAAAAACGCGGTAGAT 3011
```

```
RESULT 44
AF269294

LOCUS       AF269294                3246 bp    DNA     linear   BCT 06-DEC-2001
DEFINITION  Staphylococcus epidermidis strain SR1 clone step.1000c01 genomic
            sequence.
ACCESSION   AF269294
VERSION     AF269294.1  GI:9664615
KEYWORDS    .
SOURCE      Staphylococcus epidermidis
  ORGANISM  Staphylococcus epidermidis
            Bacteria; Firmicutes; Bacillus/Clostridium group;
            Bacillus/Staphylococcus group; Staphylococcus.
REFERENCE   1 (bases 1 to 3246)
  AUTHORS   Kimmerly,W.J., Taylor,J.D., Nelsen,A.J., Godlevski,M.M.,
            Rubio,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,l.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Transposon-mediated sequencing of the Staphylococcus epidermidis
            genome
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 3246)
  AUTHORS   Taylor,J.D., Kimmerly,W.J., Nelsen,A.J., Godlevski,M.M.,
            Rubio,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,l.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-MAY-2000) Departments of Genomic Sciences and
            Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
            Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES             Location/Qualifiers
     source          1..3246
                     /organism="Staphylococcus epidermidis"
                     /strain="SR1"
                     /db_xref="taxon:1282"
                     /clone="step.1000c01"
ORIGIN

BASE COUNT     1091 a    533 c    536 g   1086 t
ORIGIN

Query Match             24.28;  Score 16;  DB 1;  Length 3246;
Best Local Similarity   100.0%;  Pred. No. 65;
Matches   16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy        29 TAAAAAACGCGGTAGAT 44
             |||||||||||||||||
Db      2668 TAAAAAACGCGGTAGAT 2683
```

```
RESULT 45
AX144614

LOCUS       AX144614                3246 bp    DNA     linear   PAT 31-MAY-2001
DEFINITION  Sequence 3336 from Patent WO0134809.
ACCESSION   AX144614
VERSION     AX144614.1  GI:14283179
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct.
            artificial sequence.
REFERENCE   1 (bases 1 to 3246)
  AUTHORS   Kimmerly,W.J.
  TITLE     Staphylococcus epidermidis nucleic acids and proteins
  JOURNAL   Patent: WO 0134809-A 3336 17-MAY-2001;
            GLAXO GROUP LIMITED (GB)
FEATURES             Location/Qualifiers
     source          1..3246
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic nucleic acid sequence"
BASE COUNT     1091 a    533 c    536 g   1086 t
ORIGIN

Query Match             24.28;  Score 16;  DB 6;  Length 3246;
Best Local Similarity   100.0%;  Pred. No. 65;
Matches   16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy        29 TAAAAAACGCGGTAGAT 44
             |||||||||||||||||
Db      2668 TAAAAAACGCGGTAGAT 2683

Search completed: November  5, 2002, 09:20:43
Job time : 504.822 secs
```

SOURCE   Schizosaccharomyces pombe DNA.
  ORGANISM Schizosaccharomyces pombe
            Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
            Schizosaccharomycetales; Schizosaccharomycetaceae;
            Schizosaccharomyces.
REFERENCE   1 (sites)
  AUTHORS   Osaka,F., Seino,H., Seno,T. and Yamao,F.
  TITLE   A ubiquitin-conjugating enzyme in fission yeast that is essential
            for the onset of anaphase in mitosis
  JOURNAL   Mol. Cell Biol. 17 (6), 3388-3397 (1997)
  MEDLINE   97299689
REFERENCE   2 (sites)
  AUTHORS   Osaka,F., Seino,H., Seno,T. and Yamao,F.
  TITLE   An ubiquitin-conjugating enzyme in the fission yeast, essential for
            onset of anaphase in mitosis
  JOURNAL   Unpublished (1996)
REFERENCE   3 (bases 1 to 1613)
  AUTHORS   Osaka,F.
  TITLE   Direct Submission
  JOURNAL   Submitted (23-MAY-1996) Fumio Osaka, National Institute of
            Genetics, 1111 Yata, Mishima, Shizuoka 411, Japan
            (Tel:0559-81-6748)
FEATURES   Location/Qualifiers
     source   1..1613
                     /organism="Schizosaccharomyces pombe"
                     /db_xref="taxon:4896"
     gene   complement(1..35)
                     /gene="chk1"
     CDS   complement(<1..35)
                     /gene="chk1"
                     /codon_start=1
                     /product="protein kinase"
                     /protein_id="BAA20374.1"
                     /db_xref="GI:2190254"
                     /translation="MNQKLDNFPYH"
     exon   <318..412
                     /gene="ucbp4"
                     /number=1
     gene   join(318..412,461..488,539..628,779..1014,1065..1146)
                     /gene="ucbp4"
     CDS   join(318..412,461..488,539..628,779..1014,1065..1146)
                     /gene="ucbp4"
                     /note="ubiquitin-conjugating enzyme"
                     /codon_start=1
                     /product="ucbp4"
                     /protein_id="BAA20375.1"
                     /db_xref="GI:2190255"
                     /translation="MSDMQNQNPHTNSKNSSSAGMAVDGHSVTKRLRSELMSLMMSN
                     TPGISAFPDSDSNLLHWAGTTFGFSDSDYYFEGLKFKISMSFPANYPYSPPTIIFTSPMW
                     HPNVDMSGNICLDILDIKDKWSAVYNVQTILISLQSLLGEPNNASPLNAQAAELMSKDPI
                     EYKRLLMQDYKEIDEI"
     exon   318..412
                     /gene="ucbp4"
                     /number=1
     intron   413..460
                     /gene="ucbp4"
                     /number=1
     exon   461..488
                     /gene="ucbp4"
                     /number=2
     intron   489..538
                     /gene="ucbp4"
                     /number=2
     exon   539..628
                     /gene="ucbp4"
                     /number=3
     intron   629..778
                     /gene="ucbp4"
                     /number=3
     exon   779..1014
                     /gene="ucbp4"
                     /number=4
     intron   1015..1064
                     /gene="ucbp4"
                     /number=4
     exon   1065..>1146
                     /gene="ucbp4"
                     /number=4

BASE COUNT   /gene="ucbp4"
                                                    /number=5
ORIGIN   495 a   265 c   263 g   589 t   1 others

Query Match   24.2%; Score 16; DB 8; Length 1613;
Best Local Similarity 100.0%; Pred. No. 64;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   20 ACAAAAACGGTAAAAA 35
               |||||||||||||||
Db   405 ACAAAAACGGTAAAAA 420

RESULT 42
LOCUS   AF269884   3032 bp   DNA   linear   BCT 01-AUG-2000
AF269884/c
DEFINITION Staphylococcus epidermidis strain SK1 clone step.1028c04 genomic
            sequence.
ACCESSION   AF269884
VERSION   AF269884.1 GI:9623784
KEYWORDS   .
SOURCE   Staphylococcus epidermidis.
  ORGANISM Staphylococcus epidermidis
            Bacteria; Firmicutes; Bacillus/Clostridium group;
            Bacillus/Staphylococcus group; Staphylococcus.
REFERENCE   1 (bases 1 to 3032)
  AUTHORS   Kimmerly,W.J., Taylor,J.David., Nelson,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,L.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE   Transposon-mediated sequencing of the staphylococcus epidermidis
            genome
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 3032)
  AUTHORS   Taylor,J.David., Kimmerly,W.J., Nelson,A.J., Godlevski,M.M.,
            Rubino,M.A., Nelson,F.J., Rivers,P.R., Torruella-Miller,L.,
            Listenbee,S., Ashanti,C., Altshuller,G., Mamo,L., Shepherd,N.S.,
            Fuchs,R., Fleming,T., Guan,X., Du,L., Cain,D.H., Miller,G.S. and
            Furdon,P.J.
  TITLE   Direct Submission
  JOURNAL   Submitted (22-MAY-2000) Departments of Genomic Sciences and
            Bioinformatics, Genetics Directorate, Glaxo Wellcome, Inc., 5 Moore
            Drive, Research Triangle Park, North Carolina 27709-3398, USA
FEATURES   Location/Qualifiers
     source   1..3032
                     /organism="Staphylococcus epidermidis"
                     /strain="SK1"
                     /db_xref="taxon:1282"
                     /clone="step.1028c04"
BASE COUNT   939 a   600 c   422 g   1071 t
ORIGIN

Query Match   24.2%; Score 16; DB 1; Length 3032;
Best Local Similarity 100.0%; Pred. No. 65;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   29 TAAAAAACGGTAGAT 44
               |||||||||||||||
Db   3026 TAAAAAACGGTAGAT 3011

RESULT 43
LOCUS   AX145202   3032 bp   DNA   linear   PAT 41-MAY-2001
AX145202/c
DEFINITION Sequence 3924 from patent WO0134809.
ACCESSION   AX145202
VERSION   AX145202.1 GI:14283767
KEYWORDS   .
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.

FEATURES
```
                 Location/Qualifiers
     source      1..523
                 /organism="unknown"
BASE COUNT   152 a    102 c    113 g    156 t
ORIGIN

Qy      18 CGACAAAACGGTAAAA 33
           ||||||||||||||||
Db     230 CGACAAAACGGTAAAA 245
```

```
Query Match            24.28%;  Score 16;  DB 6;  Length 523;
Best Local Similarity  100.0%;  Pred. No. 63;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

RESULT 37
```
LOCUS       AR007434       535 bp  DNA   linear  PAT 04-DEC-1998
DEFINITION  Sequence 46 from patent US 5750497.
ACCESSION   AR007434
VERSION     AR007434.1  GI:3966918
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 535)
  AUTHORS   Havelund,S., Halstr.o slashed.m.J., Jonassen,I., Andersen,A.Sloth.
            and Markussen,J.
  TITLE     Acylated insulin
  JOURNAL   Patent: US 5750497-A 46 12-MAY-1998;
FEATURES             Location/Qualifiers
     source          1..535
                     /organism="unknown"
BASE COUNT   156 a    106 c    107 g    166 t
ORIGIN

Qy      18 CGACAAAACGGTAAAA 33
           ||||||||||||||||
Db     227 CGACAAAACGGTAAAA 242
```

```
Query Match            24.28%;  Score 16;  DB 6;  Length 535;
Best Local Similarity  100.0%;  Pred. No. 63;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

RESULT 38
```
LOCUS       AR007436       538 bp  DNA   linear  PAT 04-DEC-1998
DEFINITION  Sequence 49 from patent US 5750497.
ACCESSION   AR007436
VERSION     AR007436.1  GI:3966920
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 538)
  AUTHORS   Havelund,S., Halstr.o slashed.m.J., Jonassen,I., Andersen,A.Sloth.
            and Markussen,J.
  TITLE     Acylated insulin
  JOURNAL   Patent: US 5750497-A 49 12-MAY-1998;
FEATURES             Location/Qualifiers
     source          1..538
                     /organism="unknown"
BASE COUNT   156 a    107 c    107 g    168 t
ORIGIN

Qy      18 CGACAAAACGGTAAAA 33
           ||||||||||||||||
Db     227 CGACAAAACGGTAAAA 242
```

```
Query Match            24.28%;  Score 16;  DB 6;  Length 538;
Best Local Similarity  100.0%;  Pred. No. 63;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

RESULT 39
```
LOCUS       AX105687       736 bp  DNA   linear  PAT 30-APR-2001
DEFINITION  Sequence 2 from Patent WO0124811.
ACCESSION   AX105687
VERSION     AX105687.1  GI:13921709
KEYWORDS    .
SOURCE      Mus sp.
  ORGANISM  Mus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 736)
  AUTHORS   Schneider,P., Thompson,J., Cachero,T., Ambrose,C. and Rennert,P.
  TITLE     April receptor (bcma) and uses thereof
  JOURNAL   Patent: WO 0124811-A 2 12-APR-2001;
            BIOGEN, INC. (US) ; Apotech R&D S.A. (CH)
FEATURES             Location/Qualifiers
     source          1..736
                     /organism="Mus sp."
                     /db_xref="taxon:10095"
BASE COUNT   200 a    170 c    160 g    206 t
ORIGIN

Qy      18 CGACAAAACGGTAAAA 33
           ||||||||||||||||
Db     158 CGACAAAACGGTAAAA 173
```

```
Query Match            24.28%;  Score 16;  DB 6;  Length 736;
Best Local Similarity  100.0%;  Pred. No. 63;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

RESULT 40
```
LOCUS       AX079100      1168 bp  DNA   linear  PAT 22-FEB-2001
            AX079100/c
DEFINITION  Sequence 21 from Patent WO0107591.
ACCESSION   AX079100
VERSION     AX079100.1  GI:13158674
KEYWORDS    .
SOURCE      Pseudomonas sp.
  ORGANISM  Pseudomonas sp.
            Bacteria; proteobacteria.
REFERENCE   1  (bases 1 to 1168)
  AUTHORS   Fraser,C.M., Venter,C., Tuemmler,B., Hoheisel,J., Duesterhoeft,A.,
            Hilbert,H., Timmis,K.N., Moore,E., Straetz,M. and Heim,S.
  TITLE     Dna sequences which code export systems
  JOURNAL   Patent: WO 0107591-A 21 01-FEB-2001;
            THE INSTITUTE FOR GENOMIC RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1..1168
                     /organism="Pseudomonas sp."
                     /db_xref="taxon:306"
BASE COUNT   242 a    382 c    356 g    188 t
ORIGIN

Qy      14 CGTTCGACAAAACGGT 29
           ||||||||||||||||
Db     566 CGTTCGACAAAACGGT 551
```

```
Query Match            24.28%;  Score 16;  DB 6;  Length 1168;
Best Local Similarity  100.0%;  Pred. No. 64;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

RESULT 41
```
LOCUS       D85545        1613 bp  DNA   linear  PLN 14-APR-2000
DEFINITION  Yeast chk1 and ucbp4 DNA, partial and complete cds.
ACCESSION   D85545
VERSION     D85545.1  GI:2190253
KEYWORDS    UcbP4; ucbp4; protein kinase; chk1.
```

PF00005 ABC_tran"

RESULT 41

Qy   41 AAAAACGGTAGATTAC 47
         ||||||||||||||||
Db 266551 AAAAACGGTAGATTAC 266535

Query Match   25.8%; Score 17; DB 1; Length 282183;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 42
LOCUS   AR151514   61 bp   DNA   linear   PAT 08-AUG-2001
DEFINITION   Sequence 10 from patent US 6232095.
ACCESSION   AR151514
VERSION   AR151514.1 GI:15117564
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 61)
  AUTHORS   Kmiec,E.B., Holloman,W.K. and Gerhold,D.
  TITLE   Recombinant helix modification recognition proteins and uses
             thereof
  JOURNAL   Patent: US 6232095-A 10 15 MAY-2001;
FEATURES   Location/Qualifiers
     source   1..61
                     /organism="unknown"

BASE COUNT   16 a   13 c   15 g   19 t
ORIGIN

Qy   50 TCCTACGCAATTCAGCT 65
         |||||||||||||||||
Db   47 TCCTACGCAATTCAGCT 52

Query Match   24.2%; Score 16; DB 6; Length 61;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 43
LOCUS   AR007410   511 bp   DNA   linear   PAT 04-DEC-1998
DEFINITION   Sequence 40 from patent US 5750497.
ACCESSION   AR007410
VERSION   AR007410.1 GI:3966914
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 511)
  AUTHORS   Havelund,S., Halstr.o slashed.m,J., Jonassen,I., Andersen,A.Sloth.
             and Markussen,J.
  TITLE   Acylated insulin
  JOURNAL   Patent: US 5750497-A 40 12-MAY-1998;
FEATURES   Location/Qualifiers
     source   1..511
                     /organism="unknown"

BASE COUNT   153 a   98 c   104 g   156 t
ORIGIN

Qy   18 CGACAAAACGGTAAAA 33
         ||||||||||||||||
Db   227 CGACAAAACGGTAAAA 242

Query Match   24.2%; Score 16; DB 6; Length 511;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 44
AR007416

LOCUS   AR007416   523 bp   DNA   linear   PAT 04-DEC-1998
DEFINITION   Sequence 19 from patent US 5750497.
ACCESSION   AR007416
VERSION   AR007416.1 GI:3966900
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 523)
  AUTHORS   Havelund,S., Halstr.o slashed.m,J., Jonassen,I., Andersen,A.Sloth.
             and Markussen,J.
  TITLE   Acylated insulin
  JOURNAL   Patent: US 5750497-A 19 12-MAY-1998;
FEATURES   Location/Qualifiers
     source   1..523
                     /organism="unknown"

BASE COUNT   158 a   103 c   101 g   161 t
ORIGIN

Qy   18 CGACAAAACGGTAAAA 33
         ||||||||||||||||
Db   230 CGACAAAACGGTAAAA 245

Query Match   24.2%; Score 16; DB 6; Length 523;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 35
AR007426

LOCUS   AR007426   523 bp   DNA   linear   PAT 04-DEC-1998
DEFINITION   Sequence 34 from patent US 5750497.
ACCESSION   AR007426
VERSION   AR007426.1 GI:3966910
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 523)
  AUTHORS   Havelund,S., Halstr.o slashed.m,J., Jonassen,I., Andersen,A.Sloth.
             and Markussen,J.
  TITLE   Acylated insulin
  JOURNAL   Patent: US 5750497-A 34 12-MAY-1998;
FEATURES   Location/Qualifiers
     source   1..523
                     /organism="unknown"

BASE COUNT   160 a   101 c   104 g   158 t
ORIGIN

Qy   18 CGACAAAACGGTAAAA 33
         ||||||||||||||||
Db   230 CGACAAAACGGTAAAA 245

Query Match   24.2%; Score 16; DB 6; Length 523;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 36
AR007432

LOCUS   AR007432   523 bp   DNA   linear   PAT 04-DEC-1998
DEFINITION   Sequence 43 from patent US 5750497.
ACCESSION   AR007432
VERSION   AR007432.1 GI:3966916
KEYWORDS   .
SOURCE   Unknown.
  ORGANISM   Unknown.
             Unclassified.
REFERENCE   1 (bases 1 to 523)
  AUTHORS   Havelund,S., Halstr.o slashed.m,J., Jonassen,I., Andersen,A.Sloth.
             and Markussen,J.
  TITLE   Acylated insulin
  JOURNAL   Patent: US 5750497-A 43 12-MAY-1998;

```
                    HYYKMRQNYQLPIPVE"
gene                1473. .2594
                    /gene="tgt"
cds                 1473. .2594
                    /gene="tgt"
                    /EC_number="2.4.2.29"
                    /note="Cj1010, tgt, probable queuine
                    tRNA-ribosyltransferase, len: 373 aa; similar to many e.g.
                    TGT_ECOLI queuine tRNA-ribosyltransferase (EC 2.4.2.29)
                    (375 aa), fasta scores; opt: 1016 z-score: 1190.9 E(): 0,
                    41.3% identity in 373 aa overlap, 56.7% identity to
                    HP0281"
                    /codon_start=1
                    /transl_table=11
                    /product="queuine tRNA-ribosyltransferase"
                    /protein_id="CAB73266.1"
                    /db_xref="GI:6968446"
                    /translation="MEFKLKHKDGMARVCEITTAHSTFLTPIVPMPVGTVGAVKSLDAN
                    DMKNELDEAKILLANTYIHMYLRFTSKVVKDFGGLHGTIKFDRSFLTDSGGFQAFSLSKN
                    SKHFNEGIEFKSHIDGSRHLFFTPKSVLDTQVDFNSDIMMILDDLVALPATKERVKISV
                    DRTILWAKEAIIYHKNMQQNKGIGIQGONIFGILQGGTGDYEERKRCAL SILNEMPFDGLAI
                    GGLSVGEENALMYETVQNLNPYLDENRPRYLMCVGFPEDLVENVERGVDMFDCVMPTR
                    NARNGTFFTSFGKFNIKKAEFINDHEVIDSTCSCYTCRNFSRGYLNHLFKAKELTFER
                    LASLHNLHYYELARKMREAILNNSFTQFKRNFYHLRGK"
gene                2596. .3363
                    /gene="Cj1011"
cds                 2596. .3363
                    /gene="Cj1011"
                    /note="Cj1011, probable membrane protein, len: 255 aa; no
                    Hp match"
                    /codon_start=1
                    /transl_table=11
                    /product="putative membrane protein"
                    /protein_id="CAB73267.1"
                    /db_xref="GI:6968447"
                    /translation="MLEHENLKKILQAKNLETFYSEIYGQKIFYVYGLNLETWLFND
                    EKIYKLQDGEFKLSSSIEEFSNIKSILEDFKVONTHFQNLLEHKEGIILKGGVKNFY
                    KKSFVLRQKINKNLKQINLLSEAFNLLSEQAQYKKHLKILNLSISILSKNTKEHLAR
                    IDTLYTLTNAIKNEKMKNKSIYLLSISSIFLPLNLIVGFFGMNTNNLFFKDSPYGTLY
                    IFSLICILLIVGFIFYYSKKTKEFDLDEGKKAKKQTK"
trna                3416. .3492
                    /gene="tRNA Arg"
                    /note="tRNA Arg anticodon GCG, Cove score 83.29"
                    /product="tRNA-Arg"
gene                3416. .3492
                    /gene="tRNA-Arg"
                    /codon_start=1
gene                complement(3495. .3944)
                    /gene="Cj1012c"
cds                 complement(3495. .3944)
                    /gene="Cj1012c"
                    /note="Cj1012c, possible membrane protein, len: 149 aa; no
                    Hp match"
                    /codon_start=1
                    /transl_table=11
                    /product="putative membrane protein"
                    /protein_id="CAB73268.1"
                    /db_xref="GI:6968448"
                    /translation="MVIFCAALALLLEFGVIAYLIITSDGKKTIKKQKTSQKGHVAEKI
                    KKFDTDLDKMIIAASDVKLTDIELKELAKLYVQTHKLGSKTSKELDEAAKKLEFVSA
                    LAANINASAQTVSYLNKELKKIYGSYKKEIDAYEHMGLAKRKIKEDK"
gene                complement(3961. .7206)
                    /gene="Cj1013c"
cds                 complement(3961. .7206)
                    /gene="Cj1013c"
                    /note="Cj1013c, probable membrane protein, len: 1081 aa;
                    contains three domains; aa 1-90 contains three membrane
                    spanning domains; aa 90-780 is non-membrane, and aa
                    780-1081 contains ten possible membrane spanning domains.
                    Some similarity in C-teminal membrane domain to eukaryotic
                    proteins e.g. CCSA_CHLRE cytochrome C biogenesis protein
                    CCSA (353 aa), fasta scores: opt: 459 z-score: 515.4 E():
                    2.1e-21, 36.0% identity in 267 aa overlap, and to
                    prokaryotic equivalents e.g. CCMF_ECOLI cytochrome C-type
```

```
                    biogenesis protein CCMF (647 aa); BLASTP scores; E = 0.20,
                    24% identity in 293 aa overlap, 39.0% identity to HP0378
                    (domains 2+3 only)"
                    /codon_start=1
                    /transl_table=11
                    /product="putative membrane protein"
                    /protein_id="CAB73269.1"
                    /db_xref="GI:6968449"
                    /translation="MKNIIKSIGDLRVSVVLFLLFLFALFCALATFIESAYGTPTAWAMV
                    VDTFWFEYIQLLLGLLCGMFRYMFGLKLPLMIFHISFLFLVGSSAMTRYAGFEG
                    LLPIREHYQNSLLESSKTSLRISAIKDGERYSAVNDRYIGNLPFANSFKLNLGDDQ
                    AVLKYKDLILNAHYTYKENNNSDPLLVLMLSQKGSQSGVVKFEKGEVKNIEGVNPAFM
                    NDNVKAPFVRIDENLTLSSSSENLHFLSMLDGQNIDLKIGEANAKERRLYEINDISFV
                    VKAASLHAQEALEGSNRPQDESFWLWFKSAWLEVGRPMLIISTFGEPQNWKNSLLLHFK
                    DFALSNENKNLIE/ITGSNALIKLELSYRNESKEFYIPFEYNKPIMIELADQKFFISWALSY
                    EQLPFDIYLRDFVLDRYPGSMSPAFSYASEITYKNNNENDYRIFMNNVLDIGYRFYQ
                    SSYQODEKGTVLSVNKDPGKIPTYIGYFLLCLGMFMNFLNPHISRFRTLARLINKDTLK
                    HISVIFILLSFGSEKTFAQDLNSTLPVVNTINHAKALATLIVOKSADGRMVPFDILS
                    REILEKIHQSDSYKGQNSNAVMLSMLVDKWKOLEPFILMPQNQAVARDAIANILEIPS
                    AKYIAYKDFFDENRYKLQKYVENANRKNPNARGVFPIKELIKLDERANVVNLVFSGEL
                    FKFIPVQNNPNNVWLAPFSAVTTLKGDEGHIVLALIQNYFSAVENAFKDGNWIRADEG
                    LKFIKEYQEKIGYKVMFSKTKVEMEIFSNKAEIFKLAPVVLLIAGFLLLLVFSKMVV
                    PNLKISFIFKVYVLNVLAFVIHTVGLGIRAYLSGHAPWSNGYESMVYIAWALSLSGI
                    FFSRKSPIALSLTTSILSGVVLMWVAHLSEMNPQIINLVPVLNSYWLSIHVSVITASYGF
                    LGLCALLGIFTLFLMCFLKKDGRYNILNILRNIIEATRINEMAMIFGLCLLTVGNFLGA
                    IWANESWRYWSWDSKETWALVSILVYAAILHLRMIPKYCNGQFVFALWSMFAYWVIM
                    TYFGVNYFLTGLHSYAAGEAAQLPNYVVWGFALMVVLAFFARRKRNFVGKL"
gene                complement(7224. .7919)
                    /gene="livF"
cds                 complement(7224. .7919)
                    /gene="livF"
                    /note="Cj1014c, livF, probable branched-chain amino-acid
                    ABC transport system ATP-binding protein, len: 231 aa;
                    similar to e.g. LIVF_ECOLI high-affinity branched-chain
                    amino acid transport ATP-binding protein (237 aa), fasta
                    scores; opt: 686 z-score: 797.6 E(): 0, 46.4% identity in
                    233 aa overlap. No Hp ortholog. Contains PS00017
                    ATP/GTP-binding site motif A (P-loop), PS00211 ABC
                    transporters family signature, and Pfam match to entry
                    PF00005 ABC tran"
                    /codon_start=1
                    /transl_table=11
                    /product="branched-chain amino-acid ABC transport system
                    ATP-binding protein"
                    /protein_id="CAB73270.1"
                    /db_xref="GI:6968450"
                    /translation="MLIVKDLHYYGLIEAVKGIIDFKVETGHIVSLLIGSNGAGKTSTL
                    NALLNSVRTGEVNFLGYDTKRHLTHTLVQKGIAIVPECRRVFINLSVEENLKIGAFN
                    NAENYEHLREQMYKLFPKLASKKHALAGTLSGFAQMLAISRAIMSEPKLLMLDEPSL
                    GLAPKIVGEVFDIIVRLKEEGITILLVEQNAVSALKISDYAVLENGHIVMENEAKNL
                    IGDDEIRKKYLGL"
misc_feature        complement(7299. .7841)
                    /gene="livF"
                    /note="Pfam match to entry PF00005 ABC_tran, ABC
                    transporters, score 147.40, E-value 2.5e-40"
misc_feature        complement(7482. .7526)
                    /gene="livF"
                    /note="PS00211 ABC transporters family signature"
misc_feature        complement(7797. .7820)
                    /gene="livF"
                    /note="PS00017 ATP/GTP-binding site motif A (P-loop)"
gene                complement(7906. .8676)
                    /gene="livG"
cds                 complement(7906. .8676)
                    /gene="livG"
                    /note="Cj1015c, livG, probable branched-chain amino-acid
                    ABC transport system ATP-binding protein, len: 256 aa;
                    similar to e.g. LIVG_ECOLI high-affinity branched-chain
                    amino acid transport ATP-binding protein (255 aa), fasta
                    scores; opt: 680 z-score: 800.2 E(): 0, 41.2% identity in
                    250 aa overlap. No Hp ortholog. Contains PS00017
                    ATP/GTP-binding site motif A (P-loop), PS00211 ABC
                    transporters family signature, and Pfam match to entry
```

```
88329:  88428:  contig of unknown length
88429:  92543:  gap of 4104 bp in length
92544:  92612:  contig of unknown length
92613:  96701:  gap of 4069 bp in length
96702:  96801:  contig of unknown length
96802:  100597: gap of 3796 bp in length
100598: 100697: contig of unknown length
100698: 103294: gap of 2597 bp in length
103295: 103394: contig of unknown length
103395: 107762: gap of 4368 bp in length
107763: 107862: contig of unknown length
107863: 110959: gap of 3097 bp in length
110960: 111059: contig of unknown length
111060: 115331: gap of 4272 bp in length
115332: 115411: contig of unknown length
115412: 118062: gap of 2651 bp in length
118063: 118162: contig of unknown length
118163: 121065: gap of 2903 bp in length
121066: 121165: contig of unknown length
121166: 124585: gap of 3420 bp in length
124586: 124685: contig of unknown length
124685: 128030: gap of 3345 bp in length
128031: 128130: contig of unknown length
128131: 130908: gap of 2778 bp in length
130909: 131008: contig of unknown length
131009: 132555: gap of 1547 bp in length
132556: 132655: contig of unknown length
132656: 134887: gap of 2232 bp in length
134888: 134987: contig of unknown length
134988: 137795: gap of 2808 bp in length
137796: 137895: contig of unknown length
137896: 140035: gap of 2140 bp in length
140036: 140135: contig of unknown length
140136: 142552: gap of 2417 bp in length
142553: 142652: contig of unknown length
142653: 144361: gap of 1709 bp in length
144362: 144461: contig of unknown length
144462: 145897: gap of 1436 bp in length
145898: 145997: contig of unknown length
145998: 149077: gap of 3080 bp in length
149078: 149177: contig of unknown length
149178: 151266: gap of 2089 bp in length
151267: 151366: contig of unknown length
151367: 153018: gap of 1652 bp in length
153019: 153118: contig of unknown length
153119: 154617: gap of 1499 bp in length
154618: 154717: contig of unknown length
154718: 155781: gap of 1064 bp in length
155782: 155881: contig of unknown length
155882: 157463: gap of 1582 bp in length
157464: 157563: contig of unknown length
157564: 159263: gap of 1700 bp in length
159264: 159363: contig of unknown length
159364: 161101: gap of 1738 bp in length
161102: 161201: contig of unknown length
161202: 162214: gap of 1013 bp in length
162215: 162314: contig of unknown length
162315: 163548: gap of 1234 bp in length
163549: 163648: contig of unknown length
163649: 164830: gap of 1182 bp in length
164831: 164930: contig of unknown length
164931: 166488: gap of 1558 bp in length
166489: 166588: contig of unknown length
166589: 168132: gap of 1544 bp in length
168133: 168232: contig of unknown length
168233: 169752: gap of 1520 bp in length
169753: 169852: contig of unknown length
169853: 170925: gap of 1073 bp in length
170926: 171025: contig of unknown length
171026: 172095: gap of 1070 bp in length
172096: 172195: contig of unknown length
172196: 173889: gap of 1694 bp in length
173890: 173989: contig of unknown length
```

```
FEATURES
     *       173990  175362:  contig of 1373 bp in length.
                     Location/Qualifiers
             1..175362
     source
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
                     /clone="CH230-182A20"
BASE COUNT    44086 a   41768 c   40485 g   43909 t    5114 others
ORIGIN

Qy     1  TGAGATCTGGATCCGTT 17
          |||||||||||||||||
Db 155368 TGAGATCTGGATCCGTT 155352

Query Match              25.8%;  Score 17;  DB 2;  Length 175362;
Best Local Similarity   100.0%;  Prod. No. 17;
Matches 17;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

RESULT 31
CJ111168X4/c
CJ111168X4                     282183 bp    DNA        linear    BCT 08-JUL-2000
LOCUS        Campylobacter jejuni NCTC11168 complete genome; segment 4/6.
DEFINITION
ACCESSION    AL139077 AL111168
VERSION      AL139077.2  GI:6968444
KEYWORDS
SOURCE       Campylobacter jejuni.
ORGANISM     Campylobacter jejuni
             Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
             Campylobacter.
REFERENCE    1  (bases 1 to 282183)
AUTHORS      Parkhill,J., Wren,B.W., Mungall,K., Ketley,J.M., Churcher,C.,
             Basham,D., Chillingworth,T., Davies,R.M., Feltwell,T., Holroyd,S.,
             Jagels,K., Karlyshev,A., Moule,S., Pallen,M.J., Penn,C.W.,
             Quail,M., Rajandream,M.A., Rutherford,K.M., VanVliet,A.,
             Whitehead,S. and Barrell,B.G.
TITLE        The genome sequence of the food-borne pathogen Campylobacter jejuni
             reveals hypervariable sequences
JOURNAL      Nature 403 (6770), 665-668 (2000)
MEDLINE      20150912
REFERENCE    2  (bases 1 to 282183)
AUTHORS      Parkhill,J.
TITLE        Direct Submission
JOURNAL      Submitted (09-FEB-2000) Submitted on behalf of the Campylobacter
             sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
             Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
COMMENT      Notes:
             Details of C. jejuni sequencing at the Sanger Centre are available
             on the World Wide Web.
             (URL, http://www.sanger.ac.uk/Projects/C_jejuni/).
FEATURES             Location/Qualifiers
     source          1..282183
                     /organism="Campylobacter jejuni"
                     /strain="NCTC 11168"
                     /db_xref="taxon:197"
     gene            complement(43..1443)
                     /gene="Cj1009c"
     CDS             complement(43..1443)
                     /gene="Cj1009c"
                     /note="Cj1009c, unknown, len: 466 aa; 29.1% identity to
                     HP0282"
                     /codon_start=1
                     /transl_table=11
                     /product="hypothetical protein Cj1009c"
                     /protein_id="CAB73265.1"
                     /db_xref="GI:6968445"
                     /translation="MNNILLIIIDGILLAKHFLERLCFEKGLGYFFTVVCQNSEKNNLNI
                     SSEYIDLLHFEDPTSTPARLFNIMSKDFKQAPIYMQDEFETKKSYEALRSLNPNLELEIM
                     DFWGLSVNDTHANLADARMILSRRPMDFLPDIALTAQYIGLGVGEIMEVKIPASSIFA
                     YRHISSIQQKRWRIVLIYRNSKIYFVKPSFVLEPNDSILLVGDPVVLQSIFHNIRGKA
                     GQFPMFGSNVFALLDMKMNNQMDQERVILIFTTLKLFTQKSNAKRFFIHVLNPKLGYMYE
                     KLKKLSEDKEGVFFDYFNFDFKQISMWLQNNDIGLVVTDIKNFEKEKQAFFDLKIPIM
                     KVGEASFDELKFAILLSADESFELENNANVITIDLSKQLDPGVILYYYNPNSQNTLPMEE
                     YFRSLSKLVDKNIQIINKNDENPLLNLQTREDLLQFVSFQKELLNRDFARNLSTNLNR
```

NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES   Location/Qualifiers
     source   1..163156
                     /organism="Oryza sativa"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:4530"
                     /chromosome="2"
                     /clone="OJ1267_F10"
ORIGIN

BASE COUNT   45475 a   3642 c   32750 g   52239 t   50 others

Qry   19 GACAAAACGGTAAAAAA 35
           |||||||||||||||||
Db 56208 GACAAAACGGTAAAAAA 56224

RESULT 30
Query Match   25.8%; Score 17; DB 2; Length 163156;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

LOCUS   AC097220   175362 bp   DNA   linear   HTG 20-DEC-2001
DEFINITION Rattus norvegicus clone CH230-182A20, *** SEQUENCING IN PROGRESS
            ***, 51 unordered pieces.
ACCESSION   AC097220
VERSION   AC097220.3 GI:17974120
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   Norway rat.
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 175362)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia-Garcia,F.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,R., Jia,Y., Johnson,U., King,L., Korvah,J.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., Leal,B., Lewis,L.C.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Lou,seged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,F., Mawhiney,E., Mcleod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,N.,
            Morgan,N., Nguyen,L., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oraguny,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,

On Dec 20, 2001 this sequence version replaced gi:17062578.
COMMENT   Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu

            Center project name: GISM
            Center clone name: CH230-182A20

            Summary Statistics
            Assembly program: Phrap; version 0.990329First call to
            findPhrapList
            Consensus quality: 151852 bases at least Q40
            Consensus quality: 157696 bases at least Q30
            Consensus quality: 163100 bases at least Q20
            Estimated insert size: 163100 bases; sum-of-contigs estimation
            Quality coverage: 0x in Q20 bases; agarose-fp estimation
            Quality coverage: 2.4x in Q20 bases; sum-of-contigs estimation
            ----
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            *   consists of 51 contigs. The true order of the pieces
            *   is not known and their order in this sequence record is
            *   arbitrary. Gaps between the contigs are represented as
            *   runs of N, but the exact sizes of the gaps are unknown.
            *   This record will be updated with the finished sequence
            *   as soon as it is available and the accession number will
            *   be preserved.

TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 175362)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (12-OCT-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA

            Rulz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,G., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.

          1   11433: contig of 11433 bp in length
      11434   11533: gap of unknown length
      11534   23019: contig of 11486 bp in length
      23020   23119: gap of unknown length
      23120   28445: contig of 5326 bp in length
      28446   28545: gap of unknown length
      28546   35047: contig of 6502 bp in length
      35048   35147: gap of unknown length
      35148   39726: contig of 4579 bp in length
      39727   39826: gap of unknown length
      39827   45920: contig of 6094 bp in length
      45921   46020: gap of unknown length
      46021   51618: contig of 5598 bp in length
      51619   51718: gap of unknown length
      51719   57861: contig of 6143 bp in length
      57862   57961: gap of unknown length
      57962   62845: contig of 4884 bp in length
      62846   62945: gap of unknown length
      62946   67709: contig of 4764 bp in length
      67710   67809: gap of unknown length
      67810   72298: contig of 4489 bp in length
      72299   72398: gap of unknown length
      72399   77865: contig of 5467 bp in length
      77866   77965: gap of unknown length
      77966   82802: contig of 4837 bp in length
      82803   82902: gap of unknown length
      82903   88328: contig of 5426 bp in length

```
14487   14586: gap of unknown length
14587   18020: contig of 3434 bp in length
18021   18120: gap of unknown length
18121   22992: contig of 4872 bp in length
22993   23092: gap of unknown length
23093   25604: contig of 2512 bp in length
25605   25704: gap of unknown length
25705   28743: contig of 3039 bp in length
28744   28843: gap of unknown length
28844   32165: contig of 3322 bp in length
32166   32265: gap of unknown length
32266   34785: contig of 2520 bp in length
34786   34885: gap of unknown length
34886   37354: contig of 2469 bp in length
37355   37454: gap of unknown length
37455   40784: contig of 3330 bp in length
40785   40884: gap of unknown length
40885   43725: contig of 2841 bp in length
43726   43825: gap of unknown length
43826   47154: contig of 3329 bp in length
47155   47254: gap of unknown length
47255   50670: contig of 3416 bp in length
50671   50770: gap of unknown length
50771   54298: contig of 3528 bp in length
54299   54498: gap of unknown length
54499   57179: contig of 2781 bp in length
57180   57279: gap of unknown length
57280   58994: contig of 1715 bp in length
58995   59094: gap of unknown length
59095   61619: contig of 2525 bp in length
61620   61719: gap of unknown length
61720   63859: contig of 2140 bp in length
63860   63959: gap of unknown length
63960   66383: contig of 2424 bp in length
66384   66483: gap of unknown length
66484   68894: contig of 2411 bp in length
68895   68994: gap of unknown length
68995   71805: contig of 2811 bp in length
71806   71905: gap of unknown length
71906   74985: contig of 3080 bp in length
74986   75085: gap of unknown length
75086   77140: contig of 2055 bp in length
77141   77240: gap of unknown length
77241   79612: contig of 2372 bp in length
79613   79712: gap of unknown length
79713   82040: contig of 2328 bp in length
82041   82140: gap of unknown length
82141   83804: contig of 1664 bp in length
83805   83904: gap of unknown length
83905   86146: contig of 2232 bp in length
86147   86246: gap of unknown length
86247   87707: contig of 1471 bp in length
87708   87807: gap of unknown length
87808   89456: contig of 1549 bp in length
89457   89556: gap of unknown length
89557   91190: contig of 1734 bp in length
91191   91290: gap of unknown length
91291   93264: contig of 1974 bp in length
93265   93364: gap of unknown length
93365   95051: contig of 1687 bp in length
95052   95151: gap of unknown length
95152   96666: contig of 1516 bp in length
96667   96767: gap of unknown length
96768   97996: contig of 1229 bp in length
97997   98096: gap of unknown length
98097   100872: contig of 2776 bp in length
100873  100972: gap of unknown length
100973  102483: contig of 1511 bp in length
102484  102583: gap of unknown length
102584  104616: contig of 2033 bp in length
104617  104716: gap of unknown length
104717  105910: contig of 1194 bp in length
105911  106010: gap of unknown length
```

```
*  106011  107303: contig of 1293 bp in length
*  107304  107403: gap of unknown length
*  107404  109894: contig of 2491 bp in length
*  109895  109994: gap of unknown length
*  109995  111925: contig of 1931 bp in length
*  111926  112025: gap of unknown length
*  112026  114199: contig of 2174 bp in length
*  114200  114299: gap of unknown length
*  114300  115738: contig of 1439 bp in length
*  115739  115838: gap of unknown length
*  115839  117590: contig of 1752 bp in length
*  117591  117690: gap of unknown length
*  117691  119426: contig of 1736 bp in length
*  119427  119526: gap of unknown length
*  119527  120857: contig of 1331 bp in length
*  120858  120957: gap of unknown length
*  120958  122635: contig of 1678 bp in length
*  122636  122735: gap of unknown length
*  122736  123872: contig of 1137 bp in length
*  123873  123972: gap of unknown length
*  123973  125798: contig of 1826 bp in length
*  125799  125898: gap of unknown length
*  125899  126939: contig of 1041 bp in length
*  126940  127039: gap of unknown length
*  127040  128753: contig of 1714 bp in length
*  128754  128853: gap of unknown length
*  128854  130069: contig of 1216 bp in length
*  130070  130169: gap of unknown length
*  130170  131575: contig of 1406 bp in length
*  131576  131675: gap of unknown length
*  131676  133751: contig of 2076 bp in length
*  133752  133851: gap of unknown length
*  133852  135254: contig of 1403 bp in length
*  135255  135354: gap of unknown length
*  135355  136789: contig of 1435 bp in length
```

RESULT 29
AP004085

```
Query Match          25.8%;  Score 17;  DB 2;  Length 154049;
Best Local Similarity  100.0%;
Matches 17;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy       17  TCGACAAAACGGTAAA  33
             ||||||||||||||||
Db   136187  TCGACAAAACGGTAAA  136371
```

```
LOCUS       AP004085              163156 bp    DNA     linear   HTG 23-AUG-2001
DEFINITION  Oryza sativa chromosome 2 clone OJ1267_F10, *** SEQUENCING IN
            PROGRESS ***, in ordered pieces.
ACCESSION   AP004085
VERSION     AP004085.1  GI:15281364
KEYWORDS    HTG; HTGS_PHASE2.
SOURCE      Oryza sativa
  ORGANISM  Oryza sativa (cultivar:Nipponbare) DNA, clone:OJ1267_F10.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1  (bases 1 to 163156)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
            clone:OJ1267_F10
  JOURNAL   Published Only in Database (2001) In press
REFERENCE   2  (bases 1 to 163156)
  AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-AUG-2001) Takuji Sasaki, National Institute of
            Agrobiological Resources, Rice Genome Research Program; Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@nias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
COMMENT     The nucleotide sequence of this BAC clone was generated by
            combining Monsanto and RGP-Japan sequencing data.
```

RESULT 28
AC106377/c

Db   31846 AAAAAAGCGGTAGATTA 31862
            |||||||||||||||||
Qy   30 AAAAAAGCGGTAGATTA 46

Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;
Best Local Similarity 100.00%; Pred. No. 16;
Query Match   25.8%; Score 17; DB 2; Length 56766;

LOCUS   AC106377   154049 bp   DNA   linear   HTG 12-JAN-2002
DEFINITION Rattus norvegicus clone CH230-114N22, *** SEQUENCING IN PROGRESS
            ***, 71 unordered pieces.
ACCESSION   AC106377
VERSION   AC106377.1 GI:18138898
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   Rattus norvegicus
  ORGANISM Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1 (bases 1 to 154049)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., Davila,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,K., Harris,M., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Huly,K.S., Hume,J.,
            Jackson,L.E., Jacobson,D., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,N., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Liou,C., Liu,J., Liu,W.,
            Louseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., Mcleod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,J., Newton,N., Ngo,R.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,J., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,K., Tamerisa,A.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 154049)
  AUTHORS   Worley,K.C.
  TITLE   Direct Submission
  JOURNAL   Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT   Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu/
            Contact: hgsc-help@bcm.tmc.edu
            -------------- Project Information
            Center project name: GK12
            Center clone name: CH230-114N22
            ----------- Summary Statistics
            Assembly program: Phrap; version 0.990329First call to
            findPhrapList
            Consensus quality: 114472 bases at least Q40
            Consensus quality: 124676 bases at least Q30
            Consensus quality: 133577 bases at least Q20
            Estimated insert size: 117794; sum-of-contigs estimation
            Quality coverage: 0x in Q20 bases; agarose-1p estimation
            Quality coverage: 1.3x in Q20 bases; sum-of-contigs estimation
            ----
            * NOTE: Estimated insert size may differ from sequence length
            * (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 71 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            * --------------
            *   1   4766: contig of 4766 bp in length
            *   4767   4866: gap of unknown length
            *   4867   10388: contig of 5522 bp in length
            *   10389   10488: gap of unknown length
            *   10489   14486: contig of 3998 bp in length
            *   38987   39713: contig of 727 bp in length
            *   39714   39813: gap of 100 bp
            *   39814   40525: contig of 712 bp in length
            *   40526   40625: gap of 100 bp
            *   40626   41316: contig of 691 bp in length
            *   41317   41416: gap of 100 bp
            *   41417   42119: contig of 703 bp in length
            *   42120   42219: gap of 100 bp
            *   42220   42932: contig of 713 bp in length
            *   42933   43032: gap of 100 bp
            *   43033   43754: contig of 722 bp in length
            *   43755   43854: gap of 100 bp
            *   43855   44578: contig of 724 bp in length
            *   44579   44678: gap of 100 bp
            *   44679   45394: contig of 716 bp in length
            *   45395   45494: gap of 100 bp
            *   45495   46225: contig of 731 bp in length
            *   46226   46325: gap of 100 bp
            *   46326   47057: contig of 732 bp in length
            *   47058   47157: gap of 100 bp
            *   47158   47879: contig of 722 bp in length
            *   47880   47979: gap of 100 bp
            *   47980   48660: contig of 681 bp in length
            *   48661   48760: gap of 100 bp
            *   48761   49472: contig of 712 bp in length
            *   49473   49572: gap of 100 bp
            *   49573   50279: contig of 707 bp in length
            *   50280   50379: gap of 100 bp
            *   50380   51093: contig of 714 bp in length
            *   51094   51193: gap of 100 bp
            *   51194   51883: contig of 690 bp in length
            *   51884   51983: gap of 100 bp
            *   51984   52699: contig of 716 bp in length
            *   52700   52799: gap of 100 bp
            *   52800   53515: contig of 716 bp in length
            *   53516   53615: gap of 100 bp
            *   53616   54342: contig of 727 bp in length
            *   54343   54442: gap of 100 bp
            *   54443   55174: contig of 732 bp in length
            *   55175   55274: gap of 100 bp
            *   55275   55975: contig of 702 bp in length
            *   55976   56076: gap of 100 bp

JOURNAL   Unpublished

REFERENCE 2 (bases 1 to 56766)

AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
          Anderson,S., Bastien,N., Bastien,V., Boguslavsky,L., Boukhgalter,B.,
          Brown,A., Ccamarata,J., Campopiano,A., Chang,J., Chazaro,B.,
          Cheepel,Y., Colanuelo,M., Collins,S., Collymore,A., Cook,A.,
          Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
          Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
          Glinde,S., Gord,S., Goyette,M., Graham,J., Grand-Pierre,N.,
          Hagos,H., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
          Kamat,A., Karatas,A., Kells,C., Lakocque,K., Lamazares,R.,
          Landers,T., Leburzky,J., Levine,R., Liu,G., Maclean,C.,
          Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
          McKwan,P., McKernan,K., Meldrim,J., Meneus,L., Mihova,T.,
          Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
          Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J.,
          Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
          Rotta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
          Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
          Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
          Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
          Topham,K., Travers,M., Travis,N., Triqilio,J., Vassiliev,H.,
          Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,C.,
          Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE   Direct Submission

JOURNAL   Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
          Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT   All repeats were identified using RepeatMasker:
          Smit, A.F.A. & Green, P. (1996-1997)
          http://ftp.genome.washington.edu/RM/RepeatMasker.html
          ------------------- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
          ------------------- Project Information
              Center project name: L20402
              Center clone name: 207_L_14
          ----
          * NOTE: This record contains 70 individual
          * sequencing reads that have not been assembled into
          * contigs. Runs of N are used to separate the reads
          * and the order in which they appear is completely
          * arbitrary. Low-pass sequence sampling is useful for
          * identifying clones that may be gene-rich and allows
          * overlap relationships among clones to be deduced.
          * However, it should not be assumed that this clone
          * will be sequenced to completion. In the event that
          * the record is updated, the accession number will
          * be preserved.
          *
          *   720   819: gap of   100 bp
          *   820   819: contig of   719 bp in length
          *   1554   1553: contig of   734 bp in length
          *   1654   1553: gap of   100 bp
          *   2379   2478: gap of   100 bp
          *   2479   2478: contig of   725 bp in length
          *   3201   3200: contig of   722 bp in length
          *   3301   3300: gap of   100 bp
          *   4004   4003: gap of   100 bp
          *   4104   4103: contig of   703 bp in length
          *   4821   4820: contig of   717 bp in length
          *   4921   4920: gap of   100 bp
          *   5622   5621: contig of   701 bp in length
          *   5722   5721: gap of   100 bp
          *   6437   6436: contig of   715 bp in length
          *   6537   6536: gap of   100 bp
          *   7264   7263: contig of   727 bp in length
          *   7364   7263: gap of   100 bp
          *   8076   8075: contig of   712 bp in length
          *   8176   8175: gap of   100 bp
          *   8879   8878: contig of   703 bp in length
          *   8979   8978: gap of   100 bp
          *   9744   9744: contig of   756 bp in length

          *   9735   9834: gap of   100 bp
          *   9835   10521: contig of   687 bp in length
          *   10522   10621: gap of   100 bp
          *   10622   11349: contig of   728 bp in length
          *   11350   11449: gap of   100 bp
          *   11450   12161: contig of   712 bp in length
          *   12162   12261: gap of   100 bp
          *   12262   12954: contig of   693 bp in length
          *   12955   13054: gap of   100 bp
          *   13055   13733: contig of   679 bp in length
          *   13734   13833: gap of   100 bp
          *   13834   14556: contig of   723 bp in length
          *   14557   14656: gap of   100 bp
          *   14657   15375: contig of   719 bp in length
          *   15376   15475: gap of   100 bp
          *   15476   16163: contig of   688 bp in length
          *   16164   16263: gap of   100 bp
          *   16264   16980: contig of   717 bp in length
          *   16981   17080: gap of   100 bp
          *   17081   17788: contig of   708 bp in length
          *   17789   17888: gap of   100 bp
          *   17889   18607: contig of   719 bp in length
          *   18608   18707: gap of   100 bp
          *   18708   19419: contig of   712 bp in length
          *   19420   19519: gap of   100 bp
          *   19520   20208: contig of   689 bp in length
          *   20209   20308: gap of   100 bp
          *   20309   21030: contig of   722 bp in length
          *   21031   21130: gap of   100 bp
          *   21131   21848: contig of   718 bp in length
          *   21849   21948: gap of   100 bp
          *   21949   22673: contig of   725 bp in length
          *   22674   22773: gap of   100 bp
          *   22774   23479: contig of   706 bp in length
          *   23480   23579: gap of   100 bp
          *   23580   24300: contig of   721 bp in length
          *   24301   24400: gap of   100 bp
          *   24401   25104: contig of   704 bp in length
          *   25105   25204: gap of   100 bp
          *   25205   25891: contig of   687 bp in length
          *   25892   25991: gap of   100 bp
          *   25992   26704: contig of   713 bp in length
          *   26705   26805: gap of   100 bp
          *   26805   27524: contig of   720 bp in length
          *   27525   27624: gap of   100 bp
          *   27624   28334: contig of   710 bp in length
          *   28335   28434: gap of   100 bp
          *   28435   29163: contig of   729 bp in length
          *   29164   29263: gap of   100 bp
          *   29264   29975: contig of   712 bp in length
          *   29976   30075: gap of   100 bp
          *   30076   30752: contig of   677 bp in length
          *   30753   30852: gap of   100 bp
          *   30853   31580: contig of   728 bp in length
          *   31581   31680: gap of   100 bp
          *   31681   32407: contig of   727 bp in length
          *   32408   32507: gap of   100 bp
          *   32508   33224: contig of   717 bp in length
          *   33225   33324: gap of   100 bp
          *   33325   34021: contig of   697 bp in length
          *   34022   34121: gap of   100 bp
          *   34122   34826: contig of   705 bp in length
          *   34827   34926: gap of   100 bp
          *   34927   35630: contig of   704 bp in length
          *   35631   35730: gap of   100 bp
          *   35731   36442: contig of   712 bp in length
          *   36443   36542: gap of   100 bp
          *   36543   37233: contig of   691 bp in length
          *   37234   37333: gap of   100 bp
          *   37334   38059: contig of   726 bp in length
          *   38060   38159: gap of   100 bp
          *   38160   38886: contig of   727 bp in length
          *   38887   38986: gap of   100 bp

```
gene            IGPFTREKIAIVRKAIKIVEEFLRKAGLFPKVWQAFATVGEDKWVGKDRRAMGYIV
                TVRIVESEDAMTADWSRIPFEILEKISSRIISEIPEVTMVTYAVTSKPPSTIEPC"
                3457. .4848
                /gene="PAE3371"
CDS             3457. .4848
                /gene="PAE3371"
                /note="Biosynthesis of cofactors, prosthetic groups, and
                carriers; Glutathione"
                /codon_start=1
                /transl_table=11
                /product="gamma-glutamyltransferase"
                /protein_id="AAL64872.1"
                /db_xref="GI:18161601"
                /translation="MYIGTEFAIATESFLASRAGYEVYKKGGNADAVVAASAVLTYT
                LPHLGGVGGDFLAAVHKGDRAESVLGLGWAPRKIPDRPPRGIQSAVVPGYIAGLVEF
                HKRYGALPWGEVIGTALDEMRKATVHPSLAAAIERHRDLLASDPGGGRLYLALPLEPGA
                PYKIEPLLKIWGVIRONPLLFYDEIARDLEQYCYFELDDELKYRPEVKPPVFINYDGW
                VIYEAPPSLGFVVLILTVKLSQPVKSPPFSVARIKNIVAALRKAHWARIDAYLHDGEVPV
                YDILSGCVQLGEADKPEPTFGTGTTYLAAGDKELTVSAIQSLYYPPGSGFTDLKWGITFN
                NRASDFTTGINRPAPRKRPSHTLSAVVMWKEGFAWALGASAGHYRPAIYAQLIQNVIW
                YCMSPEEAILAPRFIWTGGWEVOAEEGWEAGLGVIIVKYPSRMGVAAAALRKRNNYIAA
                VADIRGDGLALGI"
tRNA            5343. .5430
                /gene="PAER20"
                /product="tRNA-Leu"
gene            5343. .5430
                /gene="PAER20"
CDS             complement(5444. .5644)
                /gene="PAE3373"
gene            complement(5444. .5644)
                /gene="PAE3373"
                /note="Hypothetical; Conserved within genome"
                /codon_start=1
                /transl_table=11
                /product="paREP2b"
                /protein_id="AAL64873.1"
                /db_xref="GI:18161602"
                /translation="MKQDGEHALVI.AALTGDCSIKEKGVSPIFTGGIYSPWQNSGEWE
                LLKWYAEVKGKQAGGCSKNYAL"
gene            5768. .6052
                /gene="PAE3375"
CDS             5768. .6052
                /gene="PAE3375"
                /note="Hypothetical; Conserved"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAL64874.1"
                /db_xref="GI:18161603"
                /translation="MRVLVKIFGPKFYFGLDTYDVITLVLDILKEGATVGDVMEFLFKRY
                PGLRQKLLRGEEIIPMHDIWNGRSIYFLQGKTALAEEGDVQIIPPFGG"
gene            complement(6049. .6993)
                /gene="PAE3376"
CDS             complement(6049. .6993)
                /gene="PAE3376"
                /note="Cellular processes; Toxin production and
                resistance"
                /codon_start=1
                /transl_table=11
                /product="beta-lactamase-like protein"
                /protein_id="AAL64875.1"
                /db_xref="GI:18161604"
                /translation="MAFFRISLPLPGMELGHVNVVVKCSPGYGLLIDAGLATYDAALS
                LRGLKSLGIKPSEIKVFITHYHADHITLAQFIAEVASPDFYIGEGEVSEVATSFED
                LARLYAEEYRRHGAPAEVAEAFLKIHPMSRYRKAFEDDVWKLPWKRVKDGELLDCGLKA
                LSTPGHTPGHTVYIYRDGVFTGDHILPKITPNISWPKKGFNPLRAYLNSLRKVYIKA
                KGYPAGHGDEISDIITRIDELVKHHEKRLGEILLAVLKEPLTTYQVAGRIKWDAGSFENF
                DVYNKIFAIGAYSHLLYLEEMGAVKRVEKDVVYWTRP"
gene            7257. .7976
                /gene="PAE3378"
CDS             7257. .7976
                /gene="PAE3378"
                /note="Regulatory functions; General"
```

```
                /codon_start=1
                /transl_table=11
                /product="sugar fermentation stimulation protein,
                conjectural"
                /protein_id="AAL64876.1"
                /db_xref="GI:18161605"
                /translation="MGIPGCGEILQLLPTPDVEGVFLKRLRNFAGVALIGGRESYVHIHD
                PGLRELLFPGAKIWARRTVRGRTEYYLAAVELDDELVLVDSSLHNKIAVWLIENGYL.
                LRGYSLEKRREPAFGRGRFDLLLKSPEGRRALVEIKGVTLEVGGRALFPDAPTSKGAHH
                MEELARAVAAGYEAHVIFLVLRKRAESMSPHWEMDRKFAEFSLLRAHKAGVKIHAVKLE
                MFKWGLRFSSELPVDLSPRDF"
gene            complement(7956. .8807)
                /gene="PAE3380"
CDS             complement(7956. .8807)
                /gene="PAE3380"
                /note="Hypothetical; Conserved"
                /codon_start=1
                /transl_table=11
                /product="conserved hypothetical protein"
                /protein_id="AAL64877.1"
                /db_xref="GI:18161606"
                /translation="MTFSIVATDGVDGVGAVSSKIVAVGAFVPHAEAAVGAVATQCYA
                NPKLCRVIISMIKQGLAAREAVERALSQDVEKELRQIGAVDIRQNAYGTTGNCEPEYA
                GHVTGHGYVALGNILAGPQVVESMARAFESQRGELVIDKLLAAILEAGDRAGGDRKKOS
                AAIIVLRPGCGYLGLTDVYVLNIRVDDHPGPVAELLRRIFKIWEILTLLQREDPLLVVLKK
                DVAGEVQEALRHIGFYKREPIGVWDEFTERAFRNNAGYENFENKIRNDDKIWGSLYKY
                LLEISRR"
gene            8846. .9430
                /gene="PAE3381"
CDS             8846. .9430
                /gene="PAE3381"
                /note="Cellular processes; Toxin production and
                resistance"
                /codon_start=1
                /transl_table=11
                /product="beta-lactamase-like protein"
                /protein_id="AAL64878.1"
                /db_xref="GI:18161607"
                /translation="MEIKNGVRLLRGSPNTLLVGNVVVDPGQPEDRAGEILSIAGRSP
                VVLTHFHADHLTATPNGAEVYAAPWGEELFVASVRARLFFTHGVVEGAVYKGSDIOL
                SGVVRPGEKIGPFEVVALPGHTFGHVGYYTEGVLYAGDALFGEKVLEKYGVPYLMDVD
                MFLSSLLDKIIELDPEFLVIGHGPLRVARGELRNW"
gene            9773. .12520
                /gene="PAE3383"
CDS             9773. .12520
                /gene="PAE3383"
                /note="Energy metabolism; Other"
                /codon_start=1
                /transl_table=11
```

Query Match   27.3%; Score 18; DB 1; Length 12645;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   17 TCGACAAAACGTAAAAA 34
               |||||||||||||||||
Db   125 TCGACAAAACGTAAAAA 142

RESULT 24
LOCUS   AX203101   66 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION Sequence 9 from Patent WO0154502.
ACCESSION   AX203101
VERSION   AX203101.1 GI:15392460
KEYWORDS   .
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0154502-A 9 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic oligonucleotide"
BASE COUNT   7 a   9 c   9 g   11 t   30 others
ORIGIN
Query Match   27.3%; Score 18; DB 6; Length 66;
  Best Local Similarity 100.0%;
  Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCAGATCTGGATCCGTTC 18
           ||||||||||||||||||
Db   1 TCAGATCTGGATCCGTTC 18

RESULT 25
LOCUS   AX203114   66 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION Sequence 22 from Patent WO0154502.
ACCESSION   AX203114
VERSION   AX203114.1 GI:15392475
KEYWORDS   .
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0154502-A 22 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT   19 a   12 c   21 g   13 t   1 others
ORIGIN
Query Match   27.3%; Score 18; DB 6; Length 66;
  Best Local Similarity 100.0%;
  Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   2 GAGATCTGGATCCGTTCG 19
           ||||||||||||||||||
Db   2 GAGATCTGGATCCGTTCG 19

RESULT 26
LOCUS   AE009927   12645 bp   DNA   linear   BCT 16-JAN-2002
DEFINITION Pyrobaculum aerophilum strain IM2 section 182 of 201 of the
            complete genome.
ACCESSION   AE009927
VERSION   AE009927.1 GI:18161597

KEYWORDS   .
SOURCE   Pyrobaculum aerophilum.
  ORGANISM Pyrobaculum aerophilum
            Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
            Thermoproteaceae; Pyrobaculum.
REFERENCE   1 (bases 1 to 12645)
  AUTHORS   Fitz-Gibbon,S.T., Ladner,H., Kim,U.J., Stetter,K.O., Simon,M.I. and
            Miller,J.H.
  TITLE   Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
            aerophilum
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 99 (2), 984-989 (2002)
   PUBMED   11792869
REFERENCE   2 (bases 1 to 12645)
  AUTHORS   Fitz-Gibbon,S.T., Ladner,H., Kim,U.-J., Stetter,K.O., Simon,M.I.
            and Miller,J.H.
  TITLE   Direct Submission
  JOURNAL   Submitted (12-DEC-2001) Microbiology and Molecular Genetics,
            University of California, Los Angeles, 405 Hilgard Ave, Los
            Angeles, CA 90095-1489, USA
FEATURES   Location/Qualifiers
     source   1..12645
                     /organism="Pyrobaculum aerophilum"
                     /strain="IM2"
                     /db_xref="taxon:13773"
     gene   73..465
                     /gene="PAE3367"
     CDS   73..465
                     /gene="PAE3367"
                     /note="hypothetical"
                     /codon_start=1
                     /trans_table=11
                     /product="hypothetical protein"
                     /protein_id="AAL64869.1"
                     /db_xref="GI:18161598"
                     /translation="MSNGVRKDAEFAVFYLKDILDKTVKIVGSSHVKRLFPDEHSAIGFIK
                     KAFTGGGCTGVITRKGPRDLTTGLVIGPAQGGKCLPKPPYTYVIOIRGFDVKLDGGLN
                     IGWLPPHHOIVVVNITFDRLLESRQIVL"
     gene   485..1720
                     /gene="PAE3368"
     CDS   485..1720
                     /gene="PAE3368"
                     /note="hypothetical; Conserved"
                     /codon_start=1
                     /trans_table=11
                     /product="conserved hypothetical protein"
                     /protein_id="AAL64870.1"
                     /db_xref="GI:18161599"
                     /translation="MELLDRALEIIRITYPLGDSCLGRLFAQMGYGLENFERGRAIKTL
                     LHMKLVDEYRKGRDVLNDLLALAKIHKFTFSRFLITSLGIIWDEAACYCCGLLDKKY
                     AKPVLPLLEGVEFRTEVGTTVPKDVVEFRESEVVKRPLLITSGESVKHEINRRIGKEHL
                     KHLTGKRVIKLRPNIVVRVDLITGKASVERNPLLEGVYLKLSRRVSQAKKFGDVKSI
                     LMDKLVIREVFGCKEHVIHVSGREDSDARMLGTIGRPLVVEVKQPIRYDGALFPFRDG
                     DIEFTPIGFTNRREVRRLKEKAKTIDIKLYRALVLSERPLTVEELNKIGELSEKTLIGY
                     TPRRIKRISPRKKVRRVRMVYELAWRQISPRVFELYVRCQGGLYVKEFIHGDGGRTYPSV
                     SEILNTHLEVLELDVLSVE"
     gene   1749..3266
                     /gene="PAE3369"
     CDS   1749..3266
                     /gene="PAE3369"
                     /note="purines, pyrimidines, nucleosides, and nucleotides;
                     Purine ribonucleotide biosynthesis"
                     /codon_start=1
                     /trans_table=11
                     /product="GMP synthetase (glutamine-hydrolysing)"
                     /protein_id="AAL64871.1"
                     /db_xref="GI:18161600"
                     /translation="MEKVLVNFGGQYAHLIARRIREVGVVAEIASPEEAVIKASKEE
                     VKAVILSGGPSSVYEPGAPDIDEGIFALSKPVLGICYGHQMIAKKLGGKVERGKGPYG
                     KTIVKILVNDPLPFDGWKPEEAVWMSHSDVEERPPPGPHVLAISENGYIAAMKGLIYG
                     VQFHPEVHHTSKGRVMFENFLRKIAKLSDVWRPEDDITRIVEEIRSRVKGDDVIVVS
                     GGVDSTVTAVLLYKAVGQRVKAVFIDHGLFRRCEPEEAASLLKSIGTDVVYIDAKERF
                     LKRLEGVADCEKRRIIGETFAEVFSIAVKQMIVNKYIAQGTLYPDVVESGAVKGAIDK
                     IKSHHNVGGLPPWFQLELIEPLREFYKDEVRRIAKALGLPEDVVYRHPFPGPGLAVRI"

RESULT 20
LOCUS   AX203096
DEFINITION Sequence 4 from Patent WO0153502.
ACCESSION   AX203096
VERSION   AX203096.1 GI:15392455
KEYWORDS
SOURCE
  ORGANISM synthetic construct.
            synthetic construct.
            artificial sequence.
REFERENCE   1 (bases 1 to 64)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 4 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..64
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="random oligonucleotide"
BASE COUNT   16 a   13 c   17 g   18 t
ORIGIN

Query Match   27.3%; Score 18; DB 6; Length 64;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   49 GTCCTACGAATTCAGCTG 66
                |||||||||||||||||||
Db   1 GTCCTACGAATTCAGCTG 18

RESULT 21
LOCUS   AX203100
DEFINITION Sequence 8 from Patent WO0153502.
ACCESSION   AX203100
VERSION   AX203100.1 GI:15392459
KEYWORDS
SOURCE   synthetic construct.
  ORGANISM synthetic construct.
            artificial sequence.
REFERENCE   1 (bases 1 to 64)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 8 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..64
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="random oligonucleotide"
BASE COUNT   22 a   11 c   16 g   15 t
ORIGIN

JOURNAL   Patent: WO 0153502-A 13 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..18
                     /organism="synthetic construct"
                     /note="primer"
                     /db_xref="taxon:32630"
BASE COUNT   4 a   5 c   4 g   5 t
ORIGIN

Qy   1 TCGAGATCTGGATCCGTTC 18
                |||||||||||||||||||
Db   1 TCGAGATCTGGATCCGTTC 18

LOCUS   AX203096   64 bp   DNA   linear   PAT 30-AUG-2001

RESULT 22
LOCUS   AX203097
DEFINITION Sequence 5 from Patent WO0153502.
ACCESSION   AX203097
VERSION   AX203097.1 GI:15392456
KEYWORDS
SOURCE
  ORGANISM synthetic construct.
            synthetic construct.
            artificial sequence.
REFERENCE   1 (bases 1 to 65)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 5 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..65
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="random oligonucleotide"
BASE COUNT   16 a   14 c   17 g   18 t
ORIGIN

Query Match   27.3%; Score 18; DB 6; Length 65;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCGAGATCTGGATCCGTTC 18
                |||||||||||||||||||
Db   1 TCGAGATCTGGATCCGTTC 18

RESULT 23
LOCUS   AX203099
DEFINITION Sequence 7 from Patent WO0153502.
ACCESSION   AX203099
VERSION   AX203099.1 GI:15392458
KEYWORDS
SOURCE   synthetic construct.
  ORGANISM synthetic construct.
            artificial sequence.
REFERENCE   1 (bases 1 to 65)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 7 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..65
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="random oligonucleotide"
BASE COUNT   19 a   13 c   18 g   15 t
ORIGIN

Query Match   27.3%; Score 18; DB 6; Length 65;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TCGAGATCTGGATCCGTTC 18
                |||||||||||||||||||
Db   1 TCGAGATCTGGATCCGTTC 18

LOCUS   AX203097   65 bp   DNA   linear   PAT 30-AUG-2001
LOCUS   AX203099   65 bp   DNA   linear   PAT 30-AUG-2001

RESULT 15
AX204112
LOCUS   AX204112   66 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION   Sequence 20 from Patent WO0153502.
ACCESSION   AX204112
VERSION   AX204112.1   GI:15392473
KEYWORDS   .
SOURCE   synthetic construct.
ORGANISM   synthetic construct
             artificial sequence.
REFERENCE   1 (bases 1 to 66)
AUTHORS   Bruce,W.B. and Niu,X.
TITLE   Novel root-preferred promoter elements and methods of use
JOURNAL   Patent: WO 0153502-A 20 26-JUL-2001;
             PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT   19 a   11 c   21 g   15 t
ORIGIN

Qy   1 TGAGATCTGGATCCGTTCG 19
          ||||||||||||||||||
Db   1 TGAGATCTGGATCCGTTCG 19

Query Match   28.8%; Score 19; DB 6; Length 66;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 16
AX204102
LOCUS   AX204102   18 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION   Sequence 10 from Patent WO0153502.
ACCESSION   AX204102
VERSION   AX204102.1   GI:15392461
KEYWORDS   .
SOURCE   synthetic construct.
ORGANISM   synthetic construct
             artificial sequence.
REFERENCE   1 (bases 1 to 18)
AUTHORS   Bruce,W.B. and Niu,X.
TITLE   Novel root-preferred promoter elements and methods of use
JOURNAL   Patent: WO 0153502-A 10 26-JUL-2001;
             PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer with BamHI site"
BASE COUNT   4 a   4 c   5 g   6 t
ORIGIN

Qy   1 TGAGATCTGGATCCGTTC 18
          ||||||||||||||||||
Db   1 TGAGATCTGGATCCGTTC 18

RESULT 17
AX203103/c
LOCUS   AX203103
DEFINITION   Sequence 11 from Patent WO0153502.

ACCESSION   AX203103
VERSION   AX203103.1   GI:15392462
KEYWORDS   .
SOURCE   synthetic construct.
ORGANISM   synthetic construct
             artificial sequence.
REFERENCE   1 (bases 1 to 18)
AUTHORS   Bruce,W.B. and Niu,X.
TITLE   Novel root-preferred promoter elements and methods of use
JOURNAL   Patent: WO 0153502-A 11 26-JUL-2001;
             PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer with EcoRI site"
BASE COUNT   5 a   5 c   4 g   4 t
ORIGIN

Qy   49 GTCCTACGGAATTCAGCTG 66
          ||||||||||||||||||
Db   18 GTCCTACGGAATTCAGCTG 1

Query Match   27.3%; Score 18; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 18
AX203104/c
LOCUS   AX203104   18 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION   Sequence 12 from Patent WO0153502.
ACCESSION   AX203104
VERSION   AX203104.1   GI:15392463
KEYWORDS   .
SOURCE   synthetic construct.
ORGANISM   synthetic construct
             artificial sequence.
REFERENCE   1 (bases 1 to 18)
AUTHORS   Bruce,W.B. and Niu,X.
TITLE   Novel root-preferred promoter elements and methods of use
JOURNAL   Patent: WO 0153502-A 12 26-JUL-2001;
             PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..18
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="primer"
BASE COUNT   6 a   5 c   4 g   3 t
ORIGIN

Qy   1 TGAGATCTGGATCCGTTC 18
          ||||||||||||||||||
Db   18 TGAGATCTGGATCCGTTC 1

Query Match   27.3%; Score 18; DB 6; Length 18;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 19
AX203105
LOCUS   AX203105   18 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION   Sequence 13 from Patent WO0153502.
ACCESSION   AX203105
VERSION   AX203105.1   GI:15392465
KEYWORDS   .
SOURCE   synthetic construct.
ORGANISM   synthetic construct
             artificial sequence.
REFERENCE   1 (bases 1 to 18)
AUTHORS   Bruce,W.B. and Niu,X.
TITLE   Novel root-preferred promoter elements and methods of use

RESULT 11
```
LOCUS       AX203113                66 bp    DNA        linear   PAT 30-AUG-2001
DEFINITION  Sequence 21 from Patent WO0153502.
ACCESSION   AX203113
VERSION     AX203113.1  GI:15392474
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE     Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 21 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES            Location/Qualifiers
  source         1..66
                 /organism="synthetic construct"
                 /db_xref="taxon:32630"
                 /note="synthetic sequences flanking a random
                 oligonucleotide"
BASE COUNT     19 a      11 c      21 g      14 t       1 others
ORIGIN

Query Match    36.4%;    Score 24;   DB 6;  Length 66;
Best Local Similarity 100.0%;  Pred. No. 0.00059;
Matches 24;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    43 ATTACCGTCCTACGAATTCAGCTG 66
         ||||||||||||||||||||||||
Db    43 ATTACCGTCCTACGAATTCAGCTG 66
```

RESULT 12
```
LOCUS       AX203115                66 bp    DNA        linear   PAT 30-AUG-2001
DEFINITION  Sequence 23 from Patent WO0153502.
ACCESSION   AX203115
VERSION     AX203115.1  GI:15392476
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE     Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 23 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES            Location/Qualifiers
  source         1..66
                 /organism="synthetic construct"
                 /db_xref="taxon:32630"
                 /note="synthetic sequences flanking a random
                 oligonucleotide"
ORIGIN

Query Match    36.4%;    Score 24;   DB 6;  Length 66;
Best Local Similarity 100.0%;  Pred. No. 0.00059;
Matches 24;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    43 ATTACCGTCCTACGAATTCAGCTG 66
         ||||||||||||||||||||||||
Db    43 ATTACCGTCCTACGAATTCAGCTG 66
```

RESULT 13
```
LOCUS       AX203106                65 bp    DNA        linear   PAT 30-AUG-2001
DEFINITION  Sequence 14 from Patent WO0153502.
ACCESSION   AX203106
VERSION     AX203106.1  GI:15392467
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 65)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE     Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 14 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES            Location/Qualifiers
  source         1..65
                 /organism="synthetic construct"
                 /db_xref="taxon:32630"
                 /note="synthetic sequences flanking a random
                 oligonucleotide"
BASE COUNT     19 a      12 c      22 g      13 t
ORIGIN

Query Match    33.3%;    Score 22;   DB 6;  Length 66;
Best Local Similarity 100.0%;  Pred. No. 0.011;
Matches 22;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    45 TACCGTCCTACGAATTCAGCTG 66
         ||||||||||||||||||||||
Db    45 TACCGTCCTACGAATTCAGCTG 66
```

RESULT 14
```
LOCUS       AX203098                62 bp    DNA        linear   PAT 30-AUG-2001
DEFINITION  Sequence 6 from Patent WO0153502.
ACCESSION   AX203098
VERSION     AX203098.1  GI:15392457
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 62)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE     Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 6 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES            Location/Qualifiers
  source         1..62
                 /organism="synthetic construct"
                 /db_xref="taxon:32630"
                 /note="random oligonucleotide"
BASE COUNT     18 a      14 c      17 g      16 t
ORIGIN

Query Match    30.3%;    Score 20;   DB 6;  Length 65;
Best Local Similarity 100.0%;  Pred. No. 0.19;
Matches 20;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 TGAGATCTGGATCCGTTCGA 20
        ||||||||||||||||||||
Db    1 TGAGATCTGGATCCGTTCGA 20
```

RESULT 15
```
BASE COUNT     18 a      11 c      18 g      15 t
ORIGIN

Query Match    28.8%;    Score 19;   DB 6;  Length 62;
Best Local Similarity 100.0%;  Pred. No. 0.79;
Matches 19;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 TGAGATCTGGATCCGTTCGG 19
```

RESULT 6
LOCUS   AX203107
DEFINITION Sequence 15 from Patent WO0153502.
ACCESSION   AX203107
VERSION   AX203107.1 GI:15392468
KEYWORDS
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 15 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT   20 a   12 c   21 g   13 t
ORIGIN

Query Match   36.4%; Score 24; DB 6; Length 66;
Best Local Similarity   100.0%; Pred. No. 0.00059;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   44 ATTACCGTCCTACGAATTCAGCTG 66
            ||||||||||||||||||||||||
Db   44 ATTACCGTCCTACGAATTCAGCTG 66

RESULT 7
LOCUS   AX203108
DEFINITION Sequence 16 from Patent WO0153502.
ACCESSION   AX203108
VERSION   AX203108.1 GI:15392469
KEYWORDS
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 16 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT   19 a   11 c   21 g   15 t
ORIGIN

Query Match   36.4%; Score 24; DB 6; Length 66;
Best Local Similarity   100.0%; Pred. No. 0.00059;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   44 ATTACCGTCCTACGAATTCAGCTG 66
            ||||||||||||||||||||||||
Db   44 ATTACCGTCCTACGAATTCAGCTG 66

RESULT 8
LOCUS   AX203109
DEFINITION Sequence 17 from Patent WO0153502.
ACCESSION   AX203109
VERSION   AX203109.1 GI:15392470
KEYWORDS

SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 17 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT   19 a   12 c   21 g   14 t
ORIGIN

Query Match   36.4%; Score 24; DB 6; Length 66;
Best Local Similarity   100.0%; Pred. No. 0.00059;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   43 ATTACCGTCCTACGAATTCAGCTG 66
            ||||||||||||||||||||||||
Db   43 ATTACCGTCCTACGAATTCAGCTG 66

RESULT 9
LOCUS   AX203110
DEFINITION Sequence 18 from Patent WO0153502.
ACCESSION   AX203110
VERSION   AX203110.1 GI:15392471
KEYWORDS
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 18 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT   19 a   12 c   21 g   14 t
ORIGIN

Query Match   36.4%; Score 24; DB 6; Length 66;
Best Local Similarity   100.0%; Pred. No. 0.00059;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   43 ATTACCGTCCTACGAATTCAGCTG 66
            ||||||||||||||||||||||||
Db   43 ATTACCGTCCTACGAATTCAGCTG 66

RESULT 10
LOCUS   AX203111
DEFINITION Sequence 19 from Patent WO0153502.
ACCESSION   AX203111
VERSION   AX203111.1 GI:15392472
KEYWORDS
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 19 26-JUL-2001;

```
Query Match                    50.0%;   Score 33;   DB 6;   Length 66;
Best Local Similarity  100.0%;   Pred. No. 1.4e-09;
Matches   33;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAA 33
           |||||||||||||||||||||||||||||||||
Db     1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAA 33

RESULT 4
AX207072
LOCUS       AX207072                26 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 25 from Patent WO0153476.
ACCESSION   AX207072
VERSION     AX207072.1  GI:15394865
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 26)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE     Novel plant promoters and methods of use
  JOURNAL   Patent: WO 0153476-A 25 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..26
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic"
BASE COUNT      12 a      4 c      6 g      4 t
ORIGIN

Query Match                    39.4%;   Score 26;   DB 6;   Length 26;
Best Local Similarity  100.0%;   Pred. No. 3.2e-05;
Matches   26;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    23 AAACGGTAAAAAAGCGGTAGATTACC 48
          ||||||||||||||||||||||||||
Db     1 AAACGGTAAAAAAGCGGTAGATTACC 26

RESULT 5
AX203093
LOCUS       AX203093                66 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 1 from Patent WO0153502.
ACCESSION   AX203093
VERSION     AX203093.1  GI:15392451
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE     Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 1 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="random oligonucleotide"
BASE COUNT      19 a     12 c     21 g     14 t
ORIGIN

Query Match                    36.4%;   Score 24;   DB 6;   Length 66;
Best Local Similarity  100.0%;   Pred. No. 0.0059;
Matches   24;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy    43 ATTACCGTCCTACGAATTCAGCTG 66
          ||||||||||||||||||||||||
Db    43 ATTACCGTCCTACGAATTCAGCTG 66
```

```
Query Match                   100.0%;   Score 66;   DB 6;   Length 66;
Best Local Similarity  100.0%;   Pred. No. 3.1e-30;
Matches   66;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGGTAGATTACCGTCCTACGAATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGGTAGATTACCGTCCTACGAATT 60

Qy    61 CAGCTG 66
           ||||||
Db    61 CAGCTG 66

RESULT 2
AX203116
LOCUS       AX203116                66 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 24 from Patent WO0153502.
ACCESSION   AX203116
VERSION     AX203116.1  GI:15392477
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE     Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 24 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT      21 a     14 c     16 g     15 t
ORIGIN

Query Match                   100.0%;   Score 66;   DB 6;   Length 66;
Best Local Similarity  100.0%;   Pred. No. 3.1e-30;
Matches   66;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy     1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGGTAGATTACCGTCCTACGAATT 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGGTAGATTACCGTCCTACGAATT 60

Qy    61 CAGCTG 66
           ||||||
Db    61 CAGCTG 66

RESULT 3
AX203095
LOCUS       AX203095                66 bp    DNA     linear   PAT 30-AUG-2001
DEFINITION  Sequence 3 from Patent WO0153502.
ACCESSION   AX203095
VERSION     AX203095.1  GI:15392454
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1  (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE     Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 3 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="random oligonucleotide"
BASE COUNT      22 a     14 c     15 g     15 t
ORIGIN
```

```
          GenCore version 5.1.3
          Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   November 5, 2002, 06:57:23 ; Search time 300.822 Seconds
                              (without alignments)
                              4591.252 Million cell updates/sec

Title:          US-09-766-113-2
Perfect score:  66
Sequence:       1 tgagatctgatccgttcga.......ccgtcctacgaattcagctg 66

Scoring table:  OLIGO_NUC
                Gapop 60.0 , Gapext 60.0

Searched:       1797656 seqs, 10463268293 residues

Word size :     0

Total number of hits satisfying chosen parameters:     3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :  GenEmbl:*
         1: qb_ba:*
         2: qb_htg:*
         3: qb_in:*
         4: qb_om:*
         5: qb_ov:*
         6: qb_pat:*
         7: qb_ph:*
         8: qb_pl:*
         9: qb_pr:*
        10: qb_ro:*
        11: qb_sts:*
        12: qb_sy:*
        13: qb_un:*
        14: qb_vi:*
        15: em_ba:*
        16: em_fun:*
        17: em_hum:*
        18: em_in:*
        19: em_mu:*
        20: em_om:*
        21: em_or:*
        22: em_ov:*
        23: em_pat:*
        24: em_ph:*
        25: em_pl:*
        26: em_ro:*
        27: em_sts:*
        28: em_un:*
        29: em_vi:*
        30: em_htg_hum:*
        31: em_htg_inv:*
        32: em_htg_other:*
        33: em_htgo_inv:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	ID	Length	N	DB ID	Description
1	66	100.0	AX203094	66	6	AX203094	Sequence
2	66	100.0	AX203116	66	6	AX203116	Sequence
3	33	50.0	AX203095	66	6	AX203095	Sequence
4	26	39.4	AX207072	26	6	AX207072	Sequence
5	24	36.4	AX203093	66	6	AX203093	Sequence
6	24	36.4	AX201107	66	6	AX201107	Sequence
7	24	36.4	AX203108	66	6	AX203108	Sequence
8	24	36.4	AX203109	66	6	AX203109	Sequence
9	24	36.4	AX203110	66	6	AX203110	Sequence
10	24	36.4	AX203111	66	6	AX203111	Sequence
11	22	33.3	AX203113	66	6	AX203113	Sequence
12	20	30.3	AX203115	65	6	AX203115	Sequence
13	19	28.8	AX203106	62	6	AX203106	Sequence
14	18	28.8	AX203098	66	6	AX203098	Sequence
15	18	27.3	AX203112	18	6	AX203112	Sequence
16	18	27.3	AX203102	18	6	AX203102	Sequence
17	18	27.3	AX201103	18	6	AX201103	Sequence
18	18	27.3	AX201104	18	6	AX201104	Sequence
19	18	27.3	AX203105	18	6	AX203105	Sequence
20	18	27.3	AX203096	64	6	AX203096	Sequence
21	18	27.3	AX203100	64	6	AX203100	Sequence
22	18	27.3	AX203097	65	6	AX203097	Sequence
23	18	27.3	AX203099	65	6	AX203099	Sequence
24	18	27.3	AX203101	66	6	AX203101	Sequence
25	18	27.3	AX203114	66	6	AX203114	Sequence
26	18	27.3	AE009927	12645	1	AE009927	Pyrobacul
27	17	25.8	AC107757	56766	2	AC107757	Mus muscu
28	17	25.8	AC106377	154049	2	AC106377	Rattus no
29	17	25.8	AP004085	163156	2	AP004085	Oryza sat
30	17	25.8	AC097220	175362	2	AC097220	Rattus no
31	17	25.8	AL139077	282183	1	CJ11168X4	Campyloba
32	16	24.2	AR151513	61	6	AR151513	Sequence
33	16	24.2	AR007430	511	6	AR007430	Sequence
34	16	24.2	AR007416	523	6	AR007416	Sequence
35	16	24.2	AR007426	523	6	AR007426	Sequence
36	16	24.2	AR007432	523	6	AR007432	Sequence
37	16	24.2	AR007434	535	6	AR007434	Sequence
38	16	24.2	AR007436	538	6	AR007436	Sequence
39	16	24.2	AX105687	736	6	AX105687	Sequence
40	16	24.2	AX079100	1168	6	AX079100	Sequence
41	16	24.2	DP5545	1613	8	DP5545	Yeast chk1
42	16	24.2	AF269884	3032	1	AF269884	Staphylor
43	16	24.2	AX145202	3032	6	AX145202	Sequence
44	16	24.2	AF269294	3246	1	AF269294	Staphylor
45	16	24.2	AX144614	3246	6	AX144614	Sequence

ALIGNMENTS

```
RESULT 1
AX203094
LOCUS       AX203094    66 bp    DNA    linear    PAT  10 AUG 2001
DEFINITION  Sequence 2 from Patent WO0153502.
ACCESSION   AX203094
VERSION     AX203094.1   GI:15392453
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE     Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 2 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:42630"
                     /note="random oligonucleotide"
BASE COUNT     21 a    14 c    16 g    15 t
ORIGIN
```

```
                  GenCore version 5.1.3
          Copyright (c) 1993 - 2002  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:       November  5, 2002, 05:28:54  ; Search time 95.8147 Seconds
                                   (without alignments)
                                   1182.662 Million cell updates/sec

Title:         US-09-766-113-2
Perfect score: 66
Sequence:      1 tgagatctggatccgttcga.........ccgtcctacgaattcagctg 66

Scoring table: OLIGO_NUC
               Gapop 60.0 , Gapext 60.0

Searched:      1736416 seqs, 858457221 residues

Word size :    0

Total number of hits satisfying chosen parameters:       3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :     N_Geneseq_032802:*
     1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
     2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
     3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
     4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
     5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
     6:  /SIDS1/gcgdata/geneseq/geneseqn-cmbl/NA1985.DAT:*
     7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
     8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
     9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
    10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
    11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
    12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
    13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
    14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
    15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
    16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
    17:  /SIDS1/gcgdata/geneseq/geneseqn-cmbl/NA1996.DAT:*
    18:  /SIDS1/gcgdata/geneseq/geneseqn-cmbl/NA1997.DAT:*
    19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
    20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
    21:  /SIDS1/gcgdata/geneseq/geneseqn-cmbl/NA2000.DAT:*
    22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
    23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
    24:  /SIDS1/gcgdata/geneseq/geneseqn-cmbl/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	66	100.0	66	22	AAS08433	A plant root-prefe
2	66	100.0	66	22	AAS08453	A plant root-prefe
3	33	50.0	66	22	AAS08434	A plant root-prefe
4	26	39.4	26	22	AAH42733	A promoter element
5	24	36.4	66	22	AAS08432	A plant root-prefe
6	24	36.4	66	22	AAS08446	A plant root-prefe
7	24	36.4	66	22	AAS08447	A plant root-prefe
8	24	36.4	66	22	AAS08448	A plant root-prefe
9	24	36.4	66	22	AAS08450	A plant root-prefe

10				66	22	AAS08452	A plant root-prefe
11				65	20	AAS08445	A plant root-prefe
12				62	19	AAS08447	A plant root-prefe
13				66	19	AAS08449	A plant root-prefe
14				18	18	AAS08441	A random oligonucl
15	c			18	18	AAS08442	A random oligonucl
16	c			18	18	AAS08443	A random oligonucl
17				18	18	AAS08444	A plant root-prefe
18				64	18	AAS08435	A plant root-prefe
19				64	18	AAS08439	A plant root-prefe
20				65	18	AAS08436	A plant root-prefe
21				65	18	AAS08438	A plant root-prefe
22				66	18	AAS08440	A random oligonucl
23				66	18	AAS08451	A plant root-prefe
24				462	23	AAS53852	Helicobacter pylor
25	c			1168	22	AAF60852	pseudomonas sp exp
26				2731	23	ABL05397	Drosophila melanog
27	c			3032	22	AAH54560	S. epidermidis gen
28				3246	22	AAH53972	S. epidermidis gen
29				4019	22	AAH54846	S. epidermidis gen
30	c			7345	23	ABL05396	Drosophila melanog
31	c			86584	21	AAF22292	BAC containing rep
32	c			36	13	AAQ24203	Transglutaminase p
33	c			44	22	AAD03192	PCR primer #2 used
34				478	7	AAN60630	Beta-urogastrone -
35				533	21	AAY91620	E.coli promoter re
36				533	22	AAF55139	Sulfometuron methy
37				539	7	AAN60632	Beta-urogastrone -
38				654	7	AAN60631	Beta-urogastrone -
39				847	22	AAI93990	Human neuroblastom
40				883	22	AAI94316	Human neuroblastom
41	c			969	21	AAZ54037	Neisseria meningit
42	c			1104	22	AAF94369	Haemophilus influe
43	c			1170	19	AAX14265	H. pylori GHRO 156
44				1457	22	AHA19747	Human nervous syst
45				1550	23	ABL02975	Drosophila melanog

ALIGNMENTS

```
RESULT 1
AAS08433
ID   AAS08433 standard; DNA; 66 BP.
XX
AC   AAS08433;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 14.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE14; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..18
FT                   /*tag= a
FT                   /label= "5'_flanking_sequence"
FT   misc_feature    19..48
FT                   /*tag= b
FT                   /note= "Randomised sequence"
FT   misc_feature    49..66
FT                   /*tag= c
FT                   /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
```

```
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Claim 5; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 21 A; 14 C; 16 G; 15 T; 0 other;

Query Match            100.0%;  Score 66;  DB 22;  Length 66;
Best Local Similarity  100.0%;  Pred. No. 4.8e-26;
Matches   66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAGCGGTAGATTACCGTCCTACGAATT 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAGCGGTAGATTACCGTCCTACGAATT 60

Qy  61 CAGCTG 66
       ||||||
Db  61 CAGCTG 66

RESULT 2
AAS08453
ID   AAS08453 standard; DNA; 66 BP.
XX
AC   AAS08453;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 21.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE21; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key              Location/Qualifiers
FT   misc_feature     1..18
FT                    /*tag=   a
FT                    /label= "5'_flanking_sequence"
FT   misc_feature     19..48
FT                    /*tag=   b
FT                    /note= "Randomised sequence"
```

```
FT   misc_feature     49..66
FT                    /*tag=   c
FT                    /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 21 A; 14 C; 16 G; 15 T; 0 other;

Query Match            100.0%;  Score 66;  DB 22;  Length 66;
Best Local Similarity  100.0%;  Pred. No. 4.8e-26;
Matches   66;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAGCGGTAGATTACCGTCCTACGAATT 60
       ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAGCGGTAGATTACCGTCCTACGAATT 60

Qy  61 CAGCTG 66
       ||||||
Db  61 CAGCTG 66

RESULT 3
AAS08434
ID   AAS08434 standard; DNA; 66 BP.
XX
AC   AAS08434;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 19.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE19; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
```

```
FH   Key              Location/Qualifiers
FT   misc_feature     1..18
FT                    /*tag-  a
FT                    /label= "5'_flanking_sequence"
FT   misc_feature     19..48
FT                    /*tag-  b
FT                    /note= "Randomised sequence"
FT   misc_feature     49..66
FT                    /*tag=  c
FT                    /label= "3'_flanking_sequence"
XX
PN   WO20015302-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI: 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Claim 5; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 66 BP: 22 A: 14 C: 15 G: 15 T; 0 other;

Query Match      50.0%;  Score 33;  DB 22;  Length 66;
Best Local Similarity   100.0%;  Pred. No. 2.6e-08;
Matches   33;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TCAGATCTGGATCCGTTCGACAAAACGGTAAAA 33
         |||||||||||||||||||||||||||||||||
Db    1 TCAGATCTGGATCCGTTCGACAAAACGGTAAAA 33

RESULT 4
AAH42733
ID   AAH42733 standard; DNA; 26 BP.
XX
AC   AAH42733;
XX
DT   01-OCT-2001  (first entry)
XX
DE   A promoter element or transcription binding site.
XX
KW   Promoter element; transcription binding site; plant promoter; SMPER;
KW   synthetic multimeric promoter element region; gene expression;
KW   insect resistance; herbicide resistance; ss.
XX
OS   Synthetic.
```

```
XX
PN   WO20015476-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02024.
XX
PR   21-JAN-2000; 2000US-0177437.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI: 2001-476118/51.
XX
PT   New plant promoters with synthetic multimeric promoter element regions,
PT   useful in plant molecular biology, particularly in regulating gene
PT   expression in plants to increase resistance against insects or
PT   herbicides -
XX
PS   Example 1; Fig 1; 67pp; English.
XX
CC   AAH42709-72 represent promoter elements or transcription binding sites.
CC   They are used to to construct synthetic multimeric promoter element
CC   regions (SMPERs). The specification describes plant promoters which
CC   comprise SMPERS. The plant promoters are useful in plant molecular
CC   biology, particularly in regulating gene expression in plants. The
CC   promoters are especially useful for transforming plants or plant cells,
CC   e.g. to increase resistance against insects or herbicides.
XX
SQ   Sequence 26 BP: 12 A: 4 C: 6 G: 4 T; 0 other;

Query Match      39.4%;  Score 26;  DB 22;  Length 26;
Best Local Similarity   100.0%;  Pred. No. 0.00016;
Matches   26;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   23 AAACGGTAAAAAACGCGGTAGATTACC 48
         ||||||||||||||| ||||||||||
Db    1 AAACGGTAAAAAAGCGGTAGATTACC 26

RESULT 5
AAS08432
ID   AAS08432 standard; DNA; 66 BP.
XX
AC   AAS08432;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 15.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE15; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key              Location/Qualifiers
FT   misc_feature     1..18
FT                    /*tag-  a
FT                    /label- "5'_flanking_sequence"
FT   misc_feature     19..48
FT                    /*tag-  b
FT                    /note- "Randomised sequence"
FT   misc_feature     49..66
FT                    /*tag=  c
FT                    /label= "3'_flanking_sequence"
XX
PN   WO20015302-A2.
XX
PD   26-JUL-2001.
```

```
XX
PF    19-JAN-2001; 2001WO-US02011.
XX
PR    21-JAN-2000; 2000US-0177473.
XX
PA    (PION-) PIONEER HI-BRED INT INC.
XX
PI    Bruce WB, Niu X;
XX
DR    WPI; 2001-442261/47.
XX
PT    Producing tissue-preferred promoter elements constructs for regulating
PT    expression of nucleotide sequences in a plant comprises identifying and
PT    isolating tissue-preferred promoter elements .
XX
PS    Claim 5; Fig 1; 45pp; English.
XX
CC    The sequence represents a plant root-preferred promoter element, RPE,
CC    isolated from a random oligonucleotide library (ROL). The invention
CC    concerns a method of identifying and isolating tissue-preferred promoter
CC    elements comprising the use of a mixture of random oligonucleotides,
CC    flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC    specific plant nuclear proteins, isolating the complexes and PCR
CC    amplifying the bound oligonucleotide. The method is used for isolating
CC    tissue-specific promoters from plants, including but not limited to
CC    root-specific promoters or root-preferred promoter elements (RPE). The
CC    RPEs are useful in the genetic manipulation of a plant when operably
CC    linked to a nucleotide sequence whose expression is to be controlled to
CC    achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC    temperature, salinity, pesticide and herbicide resistance) and biotic
CC    stress (disease resistance, resistance to attack by fungi, bacteria,
CC    viruses, insects and nematodes).
XX
SQ    Sequence 66 BP; 19 A; 12 C; 21 G; 14 T; 0 other;

Query Match          36.4%;  Score 24;  DB 22;  Length 66;
Best Local Similarity 100.0%;  Pred. No. 0.0017;
Matches   24;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    43 ATTACCGTCCTACGAATTCAGCTG 66
         ||||||||||||||||||||||||
Db    43 ATTACCGTCCTACGAATTCAGCTG 66

RESULT 6
AAS08446
ID    AAS08446 standard; DNA; 66 BP.
XX
AC    AAS08446;
XX
DT    26-SEP-2001  (first entry)
XX
DE    A plant root-preferred promoter element (RPE), RPE 18.
XX
KW    Root-preferred promoter element; RPE; abiotic stress; drought;
KW    salinity; pesticide resistance; herbicide resistance; biotic stress;
KW    disease resistance; fungal disease; bacterial disease; viral disease;
KW    insect attack; nematode attack; RPE18; random oligonucleotide library;
KW    ROL; ss.
XX
OS    Synthetic.
XX
FH    Key          Location/Qualifiers
FT    misc_feature 1..18
FT                 /tag= a
FT                 /label= "5'_flanking_sequence"
FT    misc_feature 19..48
FT                 /tag= b
FT                 /note= "Randomised sequence"
FT    misc_feature 49..66
FT                 /tag= c
FT                 /label= "3'_flanking_sequence"
XX


PN    WO200153502-A2.
XX
PD    26-JUL-2001.
XX
PF    19-JAN-2001; 2001WO-US02011.
XX
PR    21-JAN-2000; 2000US-0177473.
XX
PA    (PION-) PIONEER HI-BRED INT INC.
XX
PI    Bruce WB, Niu X;
XX
DR    WPI; 2001-442261/47.
XX
PT    Producing tissue-preferred promoter elements constructs for regulating
PT    expression of nucleotide sequences in a plant comprises identifying and
PT    isolating tissue-preferred promoter elements .
XX
PS    Example 1; Fig 1; 45pp; English.
XX
CC    The sequence represents a plant root-preferred promoter element, RPE,
CC    isolated from a random oligonucleotide library (ROL). The invention
CC    concerns a method of identifying and isolating tissue-preferred promoter
CC    elements comprising the use of a mixture of random oligonucleotides,
CC    flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC    specific plant nuclear proteins, isolating the complexes and PCR
CC    amplifying the bound oligonucleotide. The method is used for isolating
CC    tissue-specific promoters from plants, including but not limited to
CC    root-specific promoters or root-preferred promoter elements (RPE). The
CC    RPEs are useful in the genetic manipulation of a plant when operably
CC    linked to a nucleotide sequence whose expression is to be controlled to
CC    achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC    temperature, salinity, pesticide and herbicide resistance) and biotic
CC    stress (disease resistance, resistance to attack by fungi, bacteria,
CC    viruses, insects and nematodes).
XX
SQ    Sequence 66 BP; 20 A; 12 C; 21 G; 13 T; 0 other;

Query Match          36.4%;  Score 24;  DB 22;  Length 66;
Best Local Similarity 100.0%;  Pred. No. 0.0017;
Matches   24;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    43 ATTACCGTCCTACGAATTCAGCTG 66
         ||||||||||||||||||||||||
Db    43 ATTACCGTCCTACGAATTCAGCTG 66

RESULT 7
AAS08447
ID    AAS08447 standard; DNA; 66 BP.
XX
AC    AAS08447;
XX
DT    26-SEP-2001  (first entry)
XX
DE    A plant root-preferred promoter element (RPE), RPE 42.
XX
KW    Root-preferred promoter element; RPE; abiotic stress; drought;
KW    salinity; pesticide resistance; herbicide resistance; biotic stress;
KW    disease resistance; fungal disease; bacterial disease; viral disease;
KW    insect attack; nematode attack; RPE42; random oligonucleotide library;
KW    ROL; ss.
XX
OS    Synthetic.
XX
FH    Key          Location/Qualifiers
FT    misc_feature 1..18
FT                 /tag= a
FT                 /label= "5'_flanking_sequence"
FT    misc_feature 19..48
FT                 /tag= b
FT                 /note= "Randomised sequence"
FT    misc_feature 49..66
```

FT   /*tag- c
FT   /label- "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 19 A; 11 C; 21 G; 15 T; 0 other;

Query Match   36.4%; Score 24; DB 22; Length 66;
Best Local Similarity   100.0%; Pred. No. 0.0017;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   43 ATTACCGTCCTACGAATTCAGCTG 66
              ||||||||||||||||||||||||
Db   43 ATTACCGTCCTACGAATTCAGCTG 66

RESULT 8
AAS08448
ID   AAS08448 standard; DNA; 66 BP.
XX
AC   AAS08448;
XX
DT   26-SEP-2001 (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 89.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE89; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   misc_feature   1..18
FT   /*tag- a
FT   /label- "5'_flanking_sequence"
FT   misc_feature   19..48

FT   /*tag- b
FT   /note= "Randomised sequence"
FT   /*tag- c
FT   /label- "3'_flanking_sequence"
FT   misc_feature   49..66
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 19 A; 11 C; 21 G; 14 T; 1 other;

Query Match   36.4%; Score 24; DB 22; Length 66;
Best Local Similarity   100.0%; Pred. No. 0.0017;
Matches   24; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   43 ATTACCGTCCTACGAATTCAGCTG 66
              ||||||||||||||||||||||||
Db   43 ATTACCGTCCTACGAATTCAGCTG 66

RESULT 9
AAS08450
ID   AAS08450 standard; DNA; 66 BP.
XX
AC   AAS08450;
XX
DT   26-SEP-2001 (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 33.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE33; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   misc_feature   1..18

```
FT                    /*tag=   a
FT                    /label= "5'_flanking_sequence"
FT   misc_feature     19..48
FT                    /*tag=   b
FT                    /note= "Randomised sequence"
FT   misc_feature     49..66
FT                    /*tag=   c
FT                    /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB; Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 20 A; 12 C; 20 G; 13 T; 1 other;

Query Match             36.4%;  Score 24;  DB 22;  Length 66;
Best Local Similarity   100.0%;  Pred. No. 0.0017;
  Matches   24;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         43 ATTACCGTCCTACGAATTCAGCTG 66
              ||||||||||||||||||||||||
Db         43 ATTACCGTCCTACGAATTCAGCTG 66

RESULT 10
AAS08452
ID   AAS08452 standard; DNA; 66 BP.
XX
AC   AAS08452;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 73.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE73; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
```

```
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..18
FT                   /*tag=   a
FT                   /label= "5'_flanking_sequence"
FT   misc_feature    19..48
FT                   /*tag=   b
FT                   /note= "Randomised sequence"
FT   misc_feature    49..66
FT                   /*tag=   c
FT                   /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB; Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 19 A; 12 C; 22 G; 13 T; 0 other;

Query Match             33.3%;  Score 22;  DB 22;  Length 66;
Best Local Similarity   100.0%;  Pred. No. 0.021;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         45 TACCGTCCTACGAATTCAGCTG 66
              ||||||||||||||||||||||
Db         45 TACCGTCCTACGAATTCAGCTG 66

RESULT 11
AAS08445
ID   AAS08445 standard; DNA; 65 BP.
XX
AC   AAS08445;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 5.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE5; random oligonucleotide library;
```

```
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..18
FT                   /*tag= a
FT                   /label= "5'_flanking_sequence"
FT   misc_feature    19..47
FT                   /*tag= b
FT                   /note= "Randomised sequence"
FT   misc_feature    48..65
FT                   /*tag= c
FT                   /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters or root-preferred promoter elements (RPE). The
CC   RPES are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 65 BP; 18 A; 14 C; 17 G; 16 T; 0 other;

Query Match             30.3%;  Score 20;  DB 22;  Length 65;
Best Local Similarity   100.0%;  Pred. No. 0.25;
Matches   20;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TGAGATCTGGATCCGTTCGA 20
              ||||||||||||||||||||
Db          1 TGAGATCTGGATCCGTTCGA 20

RESULT 12
AAS08437
ID   AAS08437 standard; DNA; 62 BP.
XX
AC   AAS08437;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 2.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
```

```
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE2; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..18
FT                   /*tag= a
FT                   /label= "5'_flanking_sequence"
FT   misc_feature    19..44
FT                   /*tag= b
FT                   /note= "Randomised sequence"
FT   misc_feature    45..62
FT                   /*tag= c
FT                   /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Claim 5; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPES are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 62 BP; 18 A; 11 C; 18 G; 15 T; 0 other;

Query Match             28.8%;  Score 19;  DB 22;  Length 62;
Best Local Similarity   100.0%;  Pred. No. 0.85;
Matches   19;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TGAGATCTGGATCCGTTCG 19
              |||||||||||||||||||
Db          1 TGAGATCTGGATCCGTTCG 19

RESULT 13
AAS08449
ID   AAS08449 standard; DNA; 66 BP.
XX
AC   AAS08449;
XX
DT   26-SEP-2001  (first entry)
XX
```

DE   A plant root-preferred promoter element (RPE), RPE 22.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE22; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   misc_feature   1..18
FT   /*tag= a
FT   /label= "5_flanking_sequence"
FT   misc_feature   19..48
FT   /*tag= b
FT   /note= "Randomised sequence"
FT   misc_feature   49..66
FT   /*tag= c
FT   /label= "3_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US020011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI: 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 19 A; 11 C; 21 G; 15 T; 0 other;

Query Match   28.8%; Score 19; DB 22; Length 66;
Best Local Similarity   100.0%; Pred. No. 0.84;
Matches 19; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TGAGATCTGGATCCGTTCG 19
        |||||||||||||||||||
Db   1 TGAGATCTGGATCCGTTCG 19

RESULT 14
AAS08441
ID   AAS08441 standard; DNA; 18 BP.
XX
AC   AAS08441;

XX
DT   26-SEP-2001 (first entry)
XX
DE   A random oligonucleotide library, ROL, PCR primer n19808.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; random oligonucleotide library;
KW   ROL; n19808; ss; PCR primer.
XX
OS   Synthetic.
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US020011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI: 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Page 21; 45pp; English.
XX
CC   The sequence represents a PCR primer which binds to the 5' flanking
CC   region of a random oligonucleotide library (ROL) construct
CC   used to isolate tissue-specific promoter elements. The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 18 BP; 3 A; 4 C; 5 G; 6 T; 0 other;

Query Match   27.3%; Score 18; DB 22; Length 18;
Best Local Similarity   100.0%; Pred. No. 3.3;
Matches 18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TGAGATCTGGATCCGTTC 18
        ||||||||||||||||||
Db   1 TGAGATCTGGATCCGTTC 18

RESULT 15
AAS08442/c
ID   AAS08442 standard; DNA; 18 BP.
XX
AC   AAS08442;
XX
DT   26-SEP-2001 (first entry)
XX
DE   A random oligonucleotide library, ROL, PCR primer n19811.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;

```
KW   -disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; random oligonucleotide library;
KW   ROL; n19811; ss; PCR primer.
XX
OS   Synthetic.
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI: 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Page 21; 45pp; English.
XX
CC   The sequence represents a PCR primer which binds to the 3' flanking
CC   region of a random oligonucleotide library (ROL) construct. The invention
CC   used to isolate tissue-specific promoter elements. The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 18 BP: 5 A; 4 C; 5 G; 4 T; 0 other;

Query Match             27.3%;  Score 18;  DB 22;  Length 18;
Best Local Similarity   100.0%;  Pred. No. 3.3;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy         49 GTCCTACGAATTCAGCTG 66
              ||||||||||||||||||
Db         18 GTCCTACGAATTCAGCTG 1

RESULT 16
AAS08443/c
ID   AAS08443 standard; DNA; 18 BP.
XX
AC   AAS08443;
XX
DT   26-SEP-2001 (first entry)
XX
DE   A random oligonucleotide library, ROL, PCR primer n19809.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; random oligonucleotide library;
KW   ROL; n19809; ss; PCR primer.
XX
OS   Synthetic.
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI: 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Page 23; 45pp; English.
XX
CC   The sequence represents a PCR primer which binds to the 3' flanking
CC   region of a random oligonucleotide library (ROL) construct. The invention
CC   used to isolate tissue-specific promoter elements. The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 18 BP: 6 A; 5 C; 4 G; 3 T; 0 other;

Query Match             27.3%;  Score 18;  DB 22;  Length 18;
Best Local Similarity   100.0%;  Pred. No. 3.3;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TGAGATCTGGATCCGTTC 18
              ||||||||||||||||||
Db         18 TGAGATCTGGATCCGTTC 1

RESULT 17
AAS08444
ID   AAS08444 standard; DNA; 18 BP.
XX
AC   AAS08444;
XX
DT   26-SEP-2001 (first entry)
XX
DE   A random oligonucleotide library, ROL, PCR primer n19810.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; random oligonucleotide library;
KW   ROL; n19810; ss; PCR primer.
XX
OS   Synthetic.
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
```

```
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WH, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Page 23; 45pp; English.
XX
CC   The sequence represents a PCR primer which binds to the 5' flanking
CC   region of a random oligonucleotide library (ROL) construct
CC   used to isolate tissue-specific promoter elements. The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 18 BP: 4 A; 5 C; 4 G; 5 T; 0 other;

Query Match          27.3%;  Score 18;  DB 22;  Length 18;
Best Local Similarity  100.0%;  Pred. No. 3.3;
Matches  18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   49 GTCCTACGAATTCAGCTG 66
         ||||||||||||||||||
Db    1 GTCCTACGAATTCAGCTG 18

RESULT 18
AAS08435
ID   AAS08435 standard; DNA; 64 BP.
XX
AC   AAS08435;
XX
DT   26-SEP-2001 (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 29.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE29; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   misc_feature   1..18
FT                  /*tag=  a
FT                  /label= "5'_flanking_sequence"
FT   misc_feature   19..46
FT                  /*tag=  b
FT                  /note=  "Randomised sequence"
FT   misc_feature   47..64
FT                  /*tag=  c
FT                  /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
```

```
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WH, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Claim 5; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 64 BP: 16 A; 13 C; 17 G; 18 T; 0 other;

Query Match          27.3%;  Score 18;  DB 22;  Length 64;
Best Local Similarity  100.0%;  Pred. No. 2.9;
Matches  18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TGAGATCTGGATCCGTTC 18
         ||||||||||||||||||
Db    1 TGAGATCTGGATCCGTTC 18

RESULT 19
AAS08439
ID   AAS08439 standard; DNA; 64 BP.
XX
AC   AAS08439;
XX
DT   26-SEP-2001 (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 61.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE61; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   misc_feature   1..18
FT                  /*tag=  a
FT                  /label= "5'_flanking_sequence"
FT   misc_feature   19..46
FT                  /*tag=  b
FT                  /note=  "Randomised sequence"
FT   misc_feature   47..64
FT                  /*tag=  c
FT                  /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
```

```
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Claim 5; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 64 BP; 22 A; 11 C; 16 G; 15 T; 0 other;

Query Match             27.38%;  Score 18;  DB 22;  Length 64;
Best Local Similarity   100.08%;  Pred. No. 2.9;
Matches  18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TGAGATCTGGATCCGTTC 18
        ||||||||||||||||||
Db    1 TGAGATCTGGATCCGTTC 18

RESULT 20
AAS08436
ID   AAS08436 standard; DNA; 65 BP.
XX
AC   AAS08436;
XX
DT   26-SEP-2001 (first entry)
XX
DE   A plant root-preferred promoter element. (RPE), RPE 60.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE60; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..18
FT                   /*tag=  a
FT                   /label= "5'_flanking_sequence"
FT   misc_feature    19..47
FT                   /*tag=  b
FT                   /note= "Randomised sequence"
FT   misc_feature    48..65
FT                   /*tag=  c
```

```
FT                   /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Claim 5; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide, isolating the complexes and PCR
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 65 BP; 16 A; 14 C; 17 G; 18 T; 0 other;

Query Match             27.38%;  Score 18;  DB 22;  Length 65;
Best Local Similarity   100.08%;  Pred. No. 2.9;
Matches  18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    1 TGAGATCTGGATCCGTTC 18
        ||||||||||||||||||
Db    1 TGAGATCTGGATCCGTTC 18

RESULT 21
AAS08438
ID   AAS08438 standard; DNA; 65 BP.
XX
AC   AAS08438;
XX
DT   26-SEP-2001 (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 39.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE39; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    1..18
FT                   /*tag=  a
FT                   /label= "5'_flanking_sequence"
FT   misc_feature    19..47
FT                   /*tag=  b
```

```
FT   misc_feature      /note= "Randomised sequence"
FT                     48..65
FT                     /*tag= c
FT                     /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Claim 5; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 65 BP; 19 A; 13 C; 18 G; 15 T; 0 other;

Query Match             27.3%;  Score 18;  DB 22;  Length 65;
Best Local Similarity   100.0%;  Pred. No. 2.9;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 TGAGATCTGGATCCGTTC 18
         ||||||||||||||||||
Db     1 TGAGATCTGGATCCGTTC 18

RESULT 22
AAS08440
ID   AAS08440 standard; DNA; 66 BP.
XX
AC   AAS08440;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A random oligonucleotide library, ROL, sequence n19813.
XX
KW   root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; random oligonucleotide library;
KW   ROL; n19813; ss.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FT   misc_feature    1..18
FT                   /*tag= a
```

```
FT   misc_feature      /label= "5'_flanking_sequence"
FT                     19..48
FT                     /*tag= b
FT                     /note= "Randomised sequence"
FT   misc_feature      49..66
FT                     /*tag= c
FT                     /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Page 21; 45pp; English.
XX
CC   The sequence represents a random oligonucleotide library (ROL) construct
CC   used to isolate tissue-specific promoter elements. The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes);
XX
SQ   Sequence 66 BP; 7 A; 9 C; 9 G; 11 T; 30 other;

Query Match             27.3%;  Score 18;  DB 22;  Length 66;
Best Local Similarity   100.0%;  Pred. No. 2.9;
Matches   18;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy     1 TGAGATCTGGATCCGTTC 18
         ||||||||||||||||||
Db     1 TGAGATCTGGATCCGTTC 18

RESULT 23
AAS08451
ID   AAS08451 standard; DNA; 66 BP.
XX
AC   AAS08451;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 71.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE71; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
```

FH Key   Location/Qualifiers
FT misc_feature   1..18
FT   /*tag= a
FT   /label- "5'_flanking_sequence"
FT misc_feature   19..48
FT   /*tag= b
FT   /note= "Randomised sequence"
FT misc_feature   49..66
FT   /*tag= c
FT   /label- "3'_flanking_sequence"
XX
PN WO200153502-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US02011.
XX
PR 21-JAN-2000; 2000US-0177473.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Bruce WH, Niu X;
XX
DR WPI; 2001-442261/47.
XX
PT Producing tissue-preferred promoter elements constructs for regulating expression of nucleotide sequences in a plant comprises identifying and isolating tissue-preferred promoter elements -
XX
PS Example 1; Fig 1; 45pp; English.
XX
CC The sequence represents a plant root-preferred promoter element, RPE, isolated from a random oligonucleotide library (ROL). The invention concerns a method of identifying and isolating tissue-preferred promoter elements comprising the use of a mixture of random oligonucleotides, flanked by 5' and 3' flanking sequences, which hybridise to tissue-specific plant nuclear proteins, isolating the complexes and PCR amplifying the bound oligonucleotide. The method is used for isolating tissue-specific promoters from plants, including but not limited to root-specific promoters or root-preferred promoter elements (RPE). The RPEs are useful in the genetic manipulation of a plant when operably linked to a nucleotide sequence whose expression is to be controlled to achieve a desired phenotypic effect, e.g. abiotic stress (drought, temperature, salinity, pesticide and herbicide resistance) and biotic stress (disease resistance, resistance to attack by fungi, bacteria, viruses, insects and nematodes).
XX
SQ Sequence 66 BP; 19 A; 12 C; 21 G; 13 T; 1 other;
XX

Query Match   27.38; Score 18; DB 22; Length 66;
Best Local Similarity   100.0%; Pred. No. 2.9;
Matches   18; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy 2 GAGATCTGGATCCGTTCG 19
       ||||||||||||||||||
Db 2 GAGATCTGGATCCGTTCG 19

RESULT 24
AAS53852
ID AAS53852 standard; DNA; 462 BP.
XX
AC AAS53852;
XX
DT 13-FEB-2002 (first entry)
XX
DE Helicobacter pylori DNA for cellular proliferation protein #306.
XX
KW Antisense; ds; prokaryotic cellular proliferation gene; antibiotic; antibacterial; drug design.
XX
OS Helicobacter pylori.
XX

PN WO200170955-A2.
XX
PD 27-SEP-2001.
XX
PF 21-MAR-2001; 2001WO-US09180.
XX
PR 21-MAR-2000; 2000US-191078P.
PR 23-MAY-2000; 2000US-206848P.
PR 26-MAY-2000; 2000US-207727P.
PR 23-OCT-2000; 2000US-242578P.
PR 27-NOV-2000; 2000US-253625P.
PR 22-DEC-2000; 2000US-257931P.
PR 16-FEB-2001; 2001US-269308P.
XX
PA (ELIT-) ELITRA PHARM INC.
XX
PI Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ; Yamamoto RT, Xu HH;
XX
DR WPI; 2001-611495/70.
DR P-PSDB; AAU35993.
XX
PT New polynucleotides for the identification and development of antibiotics, comprise sequences of antisense nucleic acids -
XX
PS Claim 27; Seq ID No 7489; 511pp; English.
XX
CC The invention relates to antisense inhibitors of genes essential to prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an essential prokaryotic cellular proliferation protein. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 462 BP; 142 A; 78 C; 125 G; 117 T; 0 other;
XX

Query Match   24.28; Score 16; DB 23; Length 462;
Best Local Similarity   100.0%; Pred. No. 29;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   29 TAAAAAAGCGGTAGAT 44
         ||||||||||||||||
Db 192 TAAAAAAGCGGTAGAT 207

RESULT 25
AAF60852/c
ID AAF60852 standard; DNA; 1168 BP.
XX
AC AAF60852;
XX
DT 16-MAY-2001 (first entry)
XX
DE Pseudomonas sp export system associated DNA ORF11349a.
XX
KW Export system; transgenic plant; translocation; plant growth; plant development; plant yield; soil quality; phytoprotection; ds.
XX
OS Pseudomonas sp.
XX

```
PN   DE19945105-A1.
XX
PD   01-FEB-2001.
XX
PF   27-JUL-1999;   990E-1045105.
XX
PR   27-JUL-1999;   990E-1045105.
XX
PA   (TIGE ) TIGR INST GENOMIC RES.
PA   (QUIA-) QUIAGEN GMBH.
PA   (GBFB ) GES BIOTECHNOLOGISCHE FORSCHUNG MBH.
PA   (DKFZ-) DKFZ DEUT KREBSFORSCHUNGSZENTRUM.
PA   (MEDI-) MEDIZINISCHE HOCHSCHULE HANNOVER.
XX
DR   WPI; 2001-192471/20.
XX
PT   New DNA encoding a bacterial export system, useful for promoting growth
PT   and yield of plants.
XX
PS   Claim 9d; Page 24-25; 26pp; German.
XX
CC   This invention describes novel DNA sequences (1) that express products
CC   having the biological function of export systems. The invention also
CC   describes (a) recombinant expression vectors containing (1); (b)
CC   prokaryotic and eukaryotic cells transfected with (1) or
CC   the vector of (a); (c) production of export systems by culturing cells of
CC   (b); (d) (partial) expression products (II) of (1), and synthetic
CC   proteins or peptides with the same sequences; (e) mono- or poly-clonal
CC   antibodies (Ab) specific for (II); (f) hybridoma cells that produce
CC   monoclonal Ab; and (q) transgenic plants that contain cells of (b). The
CC   export systems encoded by (1) translocate, through the bacterial cell
CC   wall, endogenous or modified metabolites or endogenously or
CC   heterologously expressed proteins. (1), and their fragments, are useful:
CC   (1) for expression of recombinant export systems; (2) as probes and
CC   primers for detection, isolation and amplification of full-length cDNA
CC   sequences; and (3) for producing transgenic plants. Cells transformed
CC   with (1), also recombinant export systems or synthetic peptides or
CC   proteins with the same activities, are used to promote growth,
CC   development and yield of plants, particularly Leguminosae. Host cells
CC   that express export systems improve soil quality, are phytoprotective and
CC   improve development, growth and yield of plants.
XX
SQ   Sequence 1168 BP; 242 A; 482 C; 456 G; 188 T; 0 other;

Query Match           24.2%;  Score 16;  DB 22;   Length 1168;
Best Local Similarity   100.0%;  Pred. No. 27;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

QY        14 CGTTCGACAAAACGGT 29
             ||||||||||||||||
Db       566 CGTTCGACAAAACGGT 551
```

```
RESULT 26
AAH05497
ID   AAH05497 standard; cDNA; 2731 BP.
XX
AV   AAH05497;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 10673.
XX
KW   Drosophila; development; biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   24-MAR-2001; 2001WO-US09241.
PS
```

```
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABH61294.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions.
XX
PS   Claim 1; SEQ ID NO 10673; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABH57737-ABH72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2731 BP; 633 A; 714 C; 681 G; 703 T; 0 other;

Query Match           24.2%;  Score 16;  DB 23;   Length 2731;
Best Local Similarity   100.0%;  Pred. No. 25;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

QY        19 GACAAAACGGTAAAAA 34
             ||||||||||||||||
Db       320 GACAAAACGGTAAAAA 335
```

```
RESULT 27
AAH54560/c
ID   AAH54560 standard; DNA; 3032 BP.
XX
AC   AAH54560;
XX
DT   01-SEP-2001 (first entry)
XX
DE   S. epidermidis genomic polynucleotide sequence SEQ ID NO:3924.
XX
KW   Staphylococcus epidermidis SR1 strain; infection; diagnosis;
KW   vaccination; endocarditis; ds.
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999;  990S-0164258.
XX
PA   (GLAX ) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI; 2001-416495/33.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 8; Page 1580-1581; 2188pp; English.
```

CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG81120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4455 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 3032 BP; 939 A; 600 C; 422 G; 1071 T; 0 other;

Query Match   24.2%; Score 16; DB 22; Length 3032;
Best Local Similarity   100.0%; Pred. No. 24;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   29 TAAAAAAGCGGTAGAT 44
              ||||||||||||||||
Db   3026 TAAAAAAGCGGTAGAT 3011

RESULT 28
AAH53972
ID   AAH53972 standard; DNA; 3246 BP.
XX
AC   AAH53972;
XX
DT   04-SEP-2001 (first entry)
XX
DE   S. epidermidis genomic polynucleotide sequence SEQ ID NO:3336.
XX
KW   Staphylococcus epidermidis endocarditis; ds.
XX
OS   Staphylococcus epidermidis.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999; 99US-0164258.
XX
PA   (GLAX) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI: 2001-316495/33.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 8; Page 881-882; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG81120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the

CC   polypeptides, the polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to
CC   AAH55098 represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4455 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 3246 BP; 1091 A; 533 C; 536 G; 1086 T; 0 other;

Query Match   24.2%; Score 16; DB 22; Length 3246;
Best Local Similarity   100.0%; Pred. No. 24;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   29 TAAAAAAGCGGTAGAT 44
              ||||||||||||||||
Db   2668 TAAAAAAGCGGTAGAT 2683

RESULT 29
AAH54846/c
ID   AAH54846 standard; DNA; 4019 BP.
XX
AC   AAH54846;
XX
DT   03-SEP-2001 (first entry)
XX
OS   Staphylococcus epidermidis.
XX
KW   Staphylococcus epidermidis endocarditis; ds.
XX
DE   S. epidermidis genomic polynucleotide sequence SEQ ID NO:4210.
XX
PN   WO200134809-A2.
XX
PD   17-MAY-2001.
XX
PF   09-NOV-2000; 2000WO-US30782.
XX
PR   09-NOV-1999; 99US-0164258.
XX
PA   (GLAX) GLAXO GROUP LTD.
XX
PI   Kimmerly WJ;
XX
DR   WPI: 2001-316495/33.
XX
PT   Nucleic acids encoding polypeptides from Staphylococcus epidermidis,
PT   useful for vaccinating against infections, e.g. endocarditis -
XX
PS   Claim 8; Page 1923-1925; 2188pp; English.
XX
CC   AAH52304 to AAH53970 represent nucleic acids (I) encoding polypeptides
CC   (II), given in AAG81454 to AAG81120, from Staphylococcus epidermidis.
CC   (I) and (II) can have antibacterial activity and therefore can be used
CC   in vaccination. The nucleic acids (I) may be used to produce the
CC   S. epidermidis polypeptides (II) via the production of vectors
CC   containing them which are used to produce hosts cells which express the
CC   polypeptides. The polypeptides (II) (and/or nucleic acids) may then be
CC   used to vaccinate subjects and to raise antibodies against the bacteria.
CC   The polypeptides may also be used to assay for other inhibitors of their
CC   activity and therefore identify compounds that may be used for the
CC   treatment of S. epidermidis infections, e.g. endocarditis. AAH53971 to
CC   AAH55090 represent specifically claimed S. epidermidis genomic DNA
CC   polynucleotide sequences from the present invention. AAH55091 to

CC   AAH5509H represent oligonucleotide sequences and primers which are used
CC   in the exemplification of the present invention.
CC   N.B. The present invention specifically claims all the polynucleotide
CC   sequences given in the sequence listing of the present specification,
CC   however the sequence listing only goes up to SEQ ID NO:4454 so even
CC   though sequences are given in the disclosure for SEQ ID NO:4465 to 4472,
CC   no sequences are present for SEQ ID NO:4455 to 4464.
XX
SQ   Sequence 4019 BP; 1292 A; 738 C; 623 G; 1366 T; 0 other;

Qy   29 TAAAAAGGCGGTACAT 44
         ||||||||||||||||
Db 2617 TAAAAAGGCGGTACAT 2602

Query Match   24.2%; Score 16; DB 22; Length 4019;
Best Local Similarity 100.0%;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 40
ID   ABL05496/c
ID   ABL05496 standard; cDNA; 7445 BP.
XX
AC   ABL05496;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 10670.
XX
KW   drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
XX
PN   W020017104Z-A2.
XX
HD   27-SEP-2001.
XX
XX
PF   24-MAR-2001; 2001WO-US09241.
XX
PR   24-MAR-2000; 2000US-1916147P.
PR   11-JUL-2000; 2000US-0614150.
XX
XX
PA   (PEKE) PE CORP NY.
XX
PI   Venter JC, Adams M, Li PWD, Myers EW;
XX
XX
DR   WP1; 2001-656860/75.
DR   P-PSDB; ABH61294.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
XX
CL   Claim 1; SEQ ID NO 10670; 21pp ; Sequence listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL01840-ABL16175) and the encoded proteins
CC   sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABH57737-ABH72072).
CC   (ABH57737-ABH72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 7445 BP; 1957 A; 1658 C; 1667 G; 2063 T; 0 other;

Query Match   24.2%; Score 16; DB 23; Length 7445;
Best Local Similarity 100.0%;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   19 GACAAAACGGTAAAAA 34
         ||||||||||||||||
Db 6026 GACAAAACGGTAAAAA 6011

RESULT 31
ID   AAF22292/c
ID   AAF22292 standard; DNA; 86584 BP.
XX
AC   AAF22292;
XX
DT   20-MAR-2001 (first entry)
XX
DE   BAC containing repeats from centromeres 1-4 #15.
XX
KW   centromere; microsome; vector; ds.
XX
OS   Arabidopsis thaliana.
XX
PN   W020005532S-A2.
XX
HD   21-SEP-2000.
XX
XX
PF   17-MAR-2000; 2000WO-US07392.
XX
PR   18-MAR-1999; 99US-0125219.
PR   01-APR-1999; 99US-0127409.
PR   18-MAY-1999; 99US-0134770.
PR   13-SEP-1999; 99US-0153584.
PR   17-SEP-1999; 99US-0154603.
XX
PA   (UYCH-) UNIV CHICAGO.
XX
PI   Preuss D, Copenhaver G, Keith K;
XX
DR   WP1; 2000-587529/55.
XX
PT   Recombinant DNA construct comprising a plant centromere, useful for
PT   producing stably inherited microsomes which can serve as vectors for
PT   the construction of transgenic plant and animal cells -
XX
PS   Claim 102; Page 607-626; 1449pp; English.
XX
CC   The present invention relates to a recombinant DNA construct of a plant
CC   (Arabidopsis thaliana) centromere. The constructs are useful for
CC   producing stably inherited microsomes which can serve as vectors for
CC   the construction of transgenic plant and animal cells expressing
CC   selected proteins such as hormones, enzymes, interleukins, clotting
CC   factors, cytokines, antibodies, and growth factors.
XX
XX
SQ   Sequence 86584 BP; 26755 A; 16841 C; 17276 G; 25711 T; 1 other;

Query Match   24.2%; Score 16; DB 21; Length 86584;
Best Local Similarity 100.0%;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   11 ATCGGTTCGACAAAAC 26
         ||||||||||||||||
Db 63894 ATCGGTTCGACAAAAC 63879

RESULT 32
ID   AAQ24203/c
ID   AAQ24203 standard; DNA; 36 BP.
XX
AC   AAQ24203;
XX
DT   09-OCT-1992 (first entry)
XX
DE   Transglutaminase primer 4.
XX
KW   BTG; acyl rearrangement; deamination; PCR; signal; ss.

XX
OS   Synthetic.
XX
PN   EP481504-A.
XX
PD   22-APR-1992.
XX
PF   18-OCT-1991; 91EP-0117813.
XX
PR   19-OCT-1990; 90JP-0282566.
PA   (AJIN) AJINOMOTO CO INC.
PA   (AMAN) AMANO PHARM KK.
XX
XX
PI   Ando K, Arafuka S, Koikeda S, Matsui H, Takagi H, Washizu K;
XX
DR   WPI; 1992-1338808/17.
XX
PT   DNA fragment encoding trans:glutaminase - is inserted into
PT   vector, e.g. PnJ1053-BTG, for protein expression
XX
PS   Disclosure; Page 27; 55pp; English.
XX
CC   This DNA sequence was used in conjunction with the primer sequence
CC   given in AA024202 to amplify a portion of the BTG gene using PCR
CC   techniques. The portion of the gene amplified contained the signal,
CC   pro and structural sequences from BTG.
CC   BTG catalyses an acyl rearrangement reaction of a gamma-carboxyamide
CC   gp. of glutamine. It introduces intra- or intermolecular formation
CC   of epsilon-(gamma-Gln)-Lys cross-linking when an epsilon-amino gp.
CC   of a Lys residue acts as an acyl receptor. When water acts as an
CC   acyl acceptor the enzyme accelerates the conversion of Gln residues
CC   to Glu residues by deamination.
CC   The enzyme is used in the prodn. of gelled foods, gelled cosmetics,
CC   yogurt, gelatin, cheese etc. It is also used in the prodn. of
CC   thermally stable materials such as microcapsules and carriers of
CC   immobilized enzymes. The DNA sequence given allows the prodn. of
CC   BTG efficiently and in large quantity.
XX
SQ   Sequence 36 BP; 7 A; 15 C; 7 G; 7 T; 0 other;

Query Match   22.7%; Score 15; DB 13; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches   15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TGAGATCTGGATCCG 15
        |||||||||||||||
Db   16 TGAGATCTGGATCCG 2

RESULT 33
AAD03192/c
ID   AAD03192 standard; DNA; 44 BP.
XX
AC   AAD03192;
XX
DT   13-JUN-2001 (first entry)
XX
DE   PCR primer #2 used to isolate zea mays 70026G438 clone promoter cDNA.
KW   Corn; 70026G438 clone; promoter; morphology; physiology; growth; seed;
KW   development; nutritional enhancement; disease resistance; plant growth;
KW   environmental tolerance; chemical tolerance; pest resistance;
KW   genetic engineering; gene expression; PCR primer; ss.
OS   Zea mays.
XX
PN   WO200116307-A2.
XX
PD   08-MAR-2001.
XX
PF   30-AUG-2000; 2000WO-US23824.

PR   01-SEP-1999; 99US-0151892.
XX
PA   (RENE-) RENESSEN LLC.
XX
PI   Conner TW, Tzafrir I;
XX
DR   WPI; 2001-244405/25.
XX
PT   New plant regulatory DNA sequences, useful for selectively controlling
PT   gene expression and for modifying plants through genetic engineering to
PT   have the desired morphology, physiology, growth or nutritional
PT   enhancement
XX
PS   Example 3; Page 72; 80pp; English.
XX
CC   The present invention relates to nucleic acid sequences for regulating
CC   gene expression in plants. The promoter nucleic acid molecules are useful
CC   for controlling gene expression in plants. The promoter DNA sequences are
CC   also useful for expression of transgene products in plant embryos and
CC   plant seeds. The plant promoters are useful for modifying plants through
CC   genetic engineering to have the desired characteristics associated with
CC   plant morphology, physiology, growth and development yield, nutritional
CC   enhancement, diseases or pest resistance, or environmental or chemical
CC   tolerance.
CC   The present sequence is a PCR primer used along with an adaptor primer 2
CC   (Ar2) (AAD03171) in the nested PCR reaction to isolate Zea mays
CC   70026G438 clone promoter cDNA.
XX
SQ   Sequence 44 BP; 9 A; 18 C; 6 G; 11 T; 0 other;

Query Match   22.7%; Score 15; DB 22; Length 44;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches   15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 TGAGATCTGGATCCG 15
        |||||||||||||||
Db   17 TGAGATCTGGATCCG 3

RESULT 34
AAN60630
ID   AAN60630 standard; DNA; 478 BP.
XX
AC   AAN60630;
XX
DT   17-JUN-1991 (first entry)
XX
DE   Beta-urogastrone - beta-lactamase fusion gene from pUG2301.
KW   Beta-urogastrone gene; hormone; salivary gland; ulcers; wounds;
KW   fusion protein; beta-lactamase; ss.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..357
FT   /*tag- a
FT   CDS   1..188
FT   /product- fusion_protein
FT   /*tag- b
FT   /product- beta-lactamase
FT   misc_RNA   189..198
FT   /*tag- c
FT   /label- adaptor
FT   CDS   199..357
FT   /*tag- d
FT   /product- beta-urogastrone
XX
PN   DE3523634-A.
XX
PD   09-JAN-1986.
XX
PF   02-JUL-1985; 85DE-3523634.

```
XX
PR   02-JUL-1984; 84JP-0147691.
XX
PA   (EART-) EARTH CHEMICAL CO.
XX
PI   Aoki S, Ohgai H, Horinaka A, Hiramatsu H, Koumoto S, Nishimura A;
XX
DR   WPI; 1986-015011/04.
DR   P-PSDB; AAP60626.
XX
PN   New gene for expression of beta-urogastrone   its derivs., plasmid(s)
PT   and transformed cells contg. it.
XX
PS   Disclosure; Page 55-56; 92pp; German.
XX
CC   This sequence includes the sequence of claim 1 (bases 199-357)
CC   modified by restriction sites for coupling to the beta-lactamase
CC   gene and insertion into plasmids.  Beta-urogastrone can then
CC   easily recovered from the fusion protein expressed by transformants.
CC   The fusion protein is less easily degraded by proteases and so
CC   protects beta-urogastrone and beta-lactamase collects in the periplasm
CC   of E.coli.  It is therefore easy to collect and purify the product.
CC   Beta-urogastrone is the hormone of the salivary glands which suppresses
CC   stomach acid secretion and promotes cell growth, so is useful for
CC   treating ulcers and wounds.  Previously the product was obtd. only
CC   in small amts. from human urine.
CC   See also AAN60628, and 40-32.
XX
SQ   Sequence 478 BP; 111 A; 102 C; 125 G; 140 T; 0 other;

Qy        4 AGATCTCGGATCCGTT 17
            ||||||||||||||||
Db      467 AGATCTCGGATCCGTT 481

Query Match             22.7%;  Score 15;  DB 7;  Length 478;
Best Local Similarity   100.0%;  Pred. No. 99;
Matches   15;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;


RESULT 45
AAZ91620
ID   AAZ91620 standard; DNA; 533 BP.
XX
AC   AAZ91620;
XX
DT   17-MAY-2000  (first entry)
XX
DE   E.coli promoter region flanking sequence, DPD2087 upper.
XX
KW   Promoter region; regulatory region flanking sequence; flanking sequence;
KW   cellular insult; regulatory region identification; luciferase;
KW   luxCDABE gene complex; thermostable lux gene complex;
KW   luxCDABE gene complex; crop protection chemical;
KW   stress responsive regulatory region; ss.
XX
OS   Escherichia coli.
XX
PN   US6025141-A.
XX
PD   15-FEB-2000.
XX
PF   24-OCT-1996;  96US-0745545.
PR   24-OCT-1996;  96US-0745545.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PT   Larossa RA,  Van Dyk TK;
XX
DR   WPI; 2000-181802/16.
XX
CC   Identification of regulatory regions modulated by a cellular insult,
CC   such as crop protection chemicals -
```

```
XX
PS   Example; Column 29-30; 43pp; English.
XX
CC   This sequence is a flanking sequence for an E. coli promoter region.
CC   The invention relates to a method for identifying regulatory regions
CC   modulated by a cellular insult, comprising: (a) creating a library of
CC   gene fusions of genomic DNA fused to a promoterless, luminescent reporter
CC   gene complex selected from a gene complex encoding luciferase from
CC   Renella species, a thermostable lux gene complex, and a luxCDABE gene
CC   complex in Enteric bacteria to create fusion-containing strains;
CC   (b) culturing individual gene fusion-containing strains in liquid media;
CC   (c) contacting the fusion-containing strains at a particular growth
CC   phase with a cellular insult; and (d) analysing the fusion-containing
CC   strain for a change in luminescence, the change in luminescence
CC   indicating that the fusion-containing strain includes a regulatory region
CC   modulated by the cellular insult.  The method is useful for identifying
CC   regulatory regions affected by cellular stress such as that created by
CC   crop protection chemicals. The method can detect promoters or stress
CC   responsive regulatory regions undetectable by prior art methods.
XX
SQ   Sequence 533 BP; 166 A; 94 C; 140 G; 133 T; 0 other;

Qy       22 AAAACCGTAAAAAAG 36
            |||||||||||||||
Db       72 AAAACCGTAAAAAAG 86

Query Match             22.7%;  Score 15;  DB 21;  Length 533;
Best Local Similarity   100.0%;  Pred. No. 99;
Matches   15;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;


RESULT 36
AAF55139
ID   AAF55139 standard; DNA; 533 BP.
XX
AC   AAF55139;
XX
DT   29-MAY-2001  (first entry)
XX
DE   Sulfometuron methyl-responsive regulatory region binding sequence.
XX
KW   Escherichia coli; regulatory region; cellular insult; crop protection;
KW   sulfometuron methyl-responsive regulatory region; cellular stress;
KW   herbicide; environmental pollutant; heavy metal; oxidative damage;
KW   anaerobiosis; ss.
XX
OS   Unidentified.
XX
PN   US6194159-B1.
XX
PD   27-FEB-2001.
XX
PF   24-NOV-1999;  99US-0449081.
PR   23-OCT-1996;  96US-0735545.
XX
PA   (DUPO ) DU PONT DE NEMOURS & CO E I.
XX
PT   Larossa RA,  Van Dyk TK;
XX
DR   WPI; 2001-217925/22.
XX
CC   Novel Escherichia coli genomic segment, useful for detecting bacterial
CC   regulatory elements responsive to variety of cellular stresses,
CC   consists of sulfometuron methyl-responsive regulatory region -
PS   Claim 13; Column 29-30; 35pp; English.
XX
CC   The specification describes an Escherichia coli regulatory region
CC   modulated by a cellular insult. The regulatory region is preferably a
CC   sulfometuron methyl (SM)-responsive regulatory region, preferably
CC   bound by the sequences given in AAF55139-40. The recombinant E. coli
CC   are useful for detecting crop protection chemicals and for detecting
```

Left column (RESULT 37)

```
CC   bacterial regulatory elements responsive to a variety of cellular
CC   stresses (produced by cellular insults) such as herbicides,
CC   environmental pollutants, heavy metals, changes in temperature, changes
CC   in pH, agents producing oxidative damage, insults causing DNA damage,
CC   insults causing anaerobiosis, and biological insults such as the
CC   pathogenic life forms.
XX
SQ   Sequence 533 BP; 166 A; 94 C; 133 G; 133 T; 0 other;

Query Match      22.7%;  Score 15;  DB 22;  Length 533;
Best Local Similarity   100.0%;  Pred. No. 99;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy   22 AAAACGGTAAAAAAG 36
         |||||||||||||||
Db   72 AAAACGGTAAAAAAG 86

RESULT 37
ID   AAN60632 standard; DNA; 539 BP.
XX
AC   AAN60632;
XX
DT   17-JUN-1991 (first entry)
XX
DE   Beta-urogastrone - beta-lactamase fusion gene from pUC2701.
XX
KW   Beta-urogastrone gene; hormone; salivary gland; ulcers; wounds;
KW   fusion protein; beta-lactamase; ss.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FH
FT   CDS            1..477
FT                  /*tag= a
FT                  /product= fusion_protein
FT   CDS            1..307
FT                  /*tag= b
FT                  /product= beta-lactamase
FT   misc_RNA       308..318
FT                  /*tag= c
FT                  /label= adaptor
FT   CDS            319..477
FT                  /*tag= d
FT                  /product= beta-urogastrone
XX
PN   DE3523634-A.
XX
PD   09-JAN-1986.
XX
PF   02-JUL-1985; 85DE-3523634.
XX
PR   02-JUL-1984; 84JP-0137691.
XX
DR   WPI; 1986-015031/03.
DR   P-PSDB; AAP60628.
XX
PA   (EART-) EARTH CHEMICAL COL.
XX
PI   Aoki S, Ohgai H, Horinaka A, Hiramatsu H, Koumoto S, Nishimura A;
PI   Matsushiro S;
XX
PT   New gene for expression of beta-urogastrone - its derivs., plasmid(s)
PT   and transformed cells contg. it.
XX
PS   Disclosure; Page 59-61; 92pp; German.
XX
CC   This sequence includes the sequence of claim 1 (bases 319-477)
CC   modified by restriction sites for coupling to the beta-lactamase
CC   gene and insertion into plasmids. Beta-urogastrone can then
CC   easily recovered from the fusion protein expressed by transformants.
CC   The fusion protein is less easily degraded by proteases and so
```

Right column (RESULT 38)

```
CC   protects beta-urogastrone and beta-lactamase collects in the periplasm
CC   of E.coli. It is therefore easy to collect and purify the product.
CC   Beta-urogastrone is the hormone of the salivary glands which supresses
CC   stomach acid secretion and promotes cell growth, so is useful for
CC   treating ulcers and wounds. Previously the product was obtd. only
CC   in small amts. from human urine.
CC   See also AAN60628, and 30-32.
XX
SQ   Sequence 539 BP; 130 A; 116 C; 137 G; 156 T; 0 other;

Query Match      22.7%;  Score 15;  DB 7;  Length 539;
Best Local Similarity   100.0%;  Pred. No. 98;
Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    3 AGATCTCGATCCGTT 17
         |||||||||||||||
Db  487 AGATCTGGATCCGTT 501

RESULT 38
ID   AAN60631 standard; DNA; 654 BP.
XX
AC   AAN60631;
XX
DT   17-JUN-1991 (first entry)
XX
DE   Beta-urogastrone - beta-lactamase fusion gene from pUG2101.
XX
KW   Beta-urogastrone gene; hormone; salivary gland; ulcers; wounds;
KW   fusion protein; beta-lactamase; ss.
XX
OS   Synthetic.
XX
FH   Key            Location/Qualifiers
FH
FT   CDS            1..588
FT                  /*tag= a
FT                  /product= fusion_protein
FT   CDS            1..419
FT                  /*tag= b
FT                  /product= beta-lactamase
FT   misc_RNA       420..429
FT                  /*tag= c
FT                  /label= adaptor
FT   CDS            430..588
FT                  /*tag= d
FT                  /product= beta-urogastrone
XX
PN   DE3523634-A.
XX
PD   09-JAN-1986.
XX
PF   02-JUL-1985; 85DE-3523634.
XX
PR   02-JUL-1984; 84JP-0137691.
XX
DR   WPI; 1986-015031/03.
DR   P-PSDB; AAP60627.
XX
PA   (EART-) EARTH CHEMICAL COL.
XX
PI   Aoki S, Ohgai H, Horinaka A, Hiramatsu H, Koumoto S, Nishimura A;
PI   Matsushiro S;
XX
PT   New gene for expression of beta-urogastrone - its derivs., plasmid(s)
PT   and transformed cells contg. it.
XX
PS   Disclosure; Page 56-59; 92pp; German.
XX
CC   This sequence includes the sequence of claim 1 (bases 430-588)
CC   modified by restriction sites for coupling to the beta-lactamase
CC   gene and insertion into plasmids. Beta-urogastrone can then
CC   easily recovered from the fusion protein expressed by transformants.
```

```
CC   The fusion protein is less easily degraded by proteases and so
CC   protects beta-urogastrone and beta-lactamase collects in the periplasm
CC   of E.coli. It is therefore easy to collect and purify the product.
CC   Beta-urogastrone is the hormone of the salivary glands which supresses
CC   stomach acid secretion and promotes cell growth, so is useful for
CC   treating ulcers and wounds. Previously the product was obtd. only
CC   in small amts. from human urine.
CC   See also AAN60628, and 40-42.
XX
SQ   Sequence 654 BP; 166 A; 142 C; 163 G; 183 T; 0 other;

QY           4 AGATCTGGATCCGTT 17
               |||||||||||||||
DB         598 AGATCTGGATCCGTT 612

Query Match             22.7%;  Score 15;  DB 7;  Length 654;
Best Local Similarity   100.0%;  Pred. No. 97;
Matches    15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

RESULT 49
AA194990
ID   AA194990 standard; cDNA; 847 BP.
XX
AC   AA194990;
XX
DT   13-NOV-2001 (first entry)
XX
DE   Human neuroblastoma expressed polynucleotide SEQ ID NO 65.
XX
KW   Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
OS   Homo sapiens.
XX
PN   WO200166719-A1.
XX
PD   14-SEP-2001.
XX
PF   02-MAR-2001; 2001WO-JP01629.
XX
PR   07-MAR-2000; 2000JP-0159195.
XX
PA   (CHIB-) CHIBA PREFECTURE.
PA   (HISM ) HISAMITSU PHARM CO LTD.
XX
PI   Nakagawara A;
XX
PS   Claim 1; Page 84; 2979pp; Japanese.
XX
DR   WPI; 2001-565584/63.
XX
CC   Nucleic acids originating in gene expressed in human neuroblastoma,
CC   useful as probe or primer in diagnosing prognosis of human
CC   neuroblastoma, malignancy and susceptibility indicator or tumour marker
CC   for anti-cancer agents.
CC   The invention relates to novel genes (AA193926-AA197963) expressed in
CC   human neuroblastoma. The nucleic acids are applicable as a probe or
CC   primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC   susceptibility indicators or tumour markers for anti-cancer agents. The
CC   gene information for diagnosing prognosis is related to factors similar
CC   to that for N-myc and TrkA genes.
XX
SQ   Sequence 847 BP; 216 A; 238 C; 184 G; 186 T; 24 other;

QY           2 GAGATCTGGATCCGT 16
               ||||||||||||||
DB         212 GAGATCTGGATCCGT 226

Query Match             22.7%;  Score 15;  DB 22;  Length 847;
Best Local Similarity   100.0%;  Pred. No. 95;
Matches    15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;


RESULT 40
AA194316
ID   AA194316 standard; cDNA; 883 BP.
XX
AC   AA194316;
XX
DT   13-NOV-2001 (first entry)
XX
DE   Human neuroblastoma expressed polynucleotide SEQ ID NO 391.
XX
KW   Human; neuroblastoma; malignancy; cancer; tumour marker; N-myc; TrkA; ss.
XX
OS   Homo sapiens.
XX
PN   WO200166719-A1.
XX
PD   14-SEP-2001.
XX
PF   02-MAR-2001; 2001WO-JP01629.
XX
PR   07-MAR-2000; 2000JP-0159195.
XX
PA   (CHIB-) CHIBA PREFECTURE.
PA   (HISM ) HISAMITSU PHARM CO LTD.
XX
PI   Nakagawara A;
XX
PS   Claim 1; Page 330; 2979pp; Japanese.
XX
DR   WPI; 2001-565584/63.
XX
CC   Nucleic acids originating in gene expressed in human neuroblastoma,
CC   useful as probe or primer in diagnosing prognosis of human
CC   neuroblastoma, malignancy and susceptibility indicator or tumour marker
CC   for anti-cancer agents.
CC   The invention relates to novel genes (AA193926-AA197963) expressed in
CC   human neuroblastoma. The nucleic acids are applicable as a probe or
CC   primer in diagnosing the prognosis of human neuroblastoma, malignancy and
CC   susceptibility indicators or tumour markers for anti-cancer agents. The
CC   gene information for diagnosing prognosis is related to factors similar
CC   to that for N-myc and TrkA genes.
XX
SQ   Sequence 883 BP; 235 A; 232 C; 187 G; 178 T; 51 other;

QY           2 GAGATCTGGATCCGT 16
               ||||||||||||||
DB         189 GAGATCTGGATCCGT 203

Query Match             22.7%;  Score 15;  DB 22;  Length 883;
Best Local Similarity   100.0%;  Pred. No. 94;
Matches    15;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

RESULT 41
AAZ54037/C
ID   AAZ54037 standard; DNA; 969 BP.
XX
AC   AAZ54037;
XX
DT   21-MAR-2000 (first entry)
XX
DE   Neisseria meningitidis ORF 629 partial DNA sequence SEQ ID NO:2023.
XX
KW   Neisseria meningitidis; Neisseria gonorrhoea; antigen; vaccine;
KW   antigenic; diagnosis; immunogenic; infection; meningitis; septicaemia;
KW   antibacterial; gene therapy; ds.
XX
OS   Neisseria meningitidis.
XX
PN   WO9957280-A2.
```

PD   11-NOV-1999.
XX
PF   30-APR-1999;
XX
PR   99WO-US09346.
XX
PR   01-MAY-1998;   98US-0083758.
PR   31-JUL-1998;   98US-0094869.
PR   02-SEP-1998;   98US-0098994.
PR   02-SEP-1998;   98US-0099062.
PR   09-OCT-1998;   98US-0103749.
PR   09-OCT-1998;   98US-0103794.
PR   09-OCT-1998;   98US-0103796.
PR   25-FEB-1999;   99US-0121528.
XX
PA   (CHIR) CHIRON CORP.
PA   (GENO-) INST GENOMIC RES.
XX
PI   Fraser C, Galeotti C, Grandi G, Hickey E, Masignani V, Mora M;
PI   Petersen J, Pizza M, Rappuoli R, Ratti G, Scalato E, Scarselli M;
PI   Tettelin H, Venter JC;
XX
DR   WPI: 2000-062150/05.
DR   P-PSDB: AAY75275.
XX
PT   Novel Neisserial polypeptides predicted to be useful antigens for
PT   vaccines and diagnostics -
XX
PS   Claim 7; Page 1005; 1453pp; English.
XX
CC   AAZ53015 to AAZ54536, AAZ54577 to AAZ54615, and AAY74253 to AAY75941
CC   represent novel Neisseria meningitis and N. gonorrheae polynucleotides
CC   and polypeptides. AAZ54537 to AAZ54576 and AAZ54616 to AAZ54473 represent
CC   PCR primers used in the exemplification of the present invention. The
CC   polynucleotides, the polynucleotides, antibodies and compositions of
CC   the invention can be used as vaccines, as diagnostic reagents, and as
CC   immunogenic compositions. The polypeptides can be used in the
CC   manufacture of medicaments for treating or preventing infection due to
CC   Neisserial bacteria (e.g. meningitis and septicaemia), to detect the
CC   presence of Neisseria bacteria, or to raise antibodies. They may also
CC   be used to screen for agonists or antagonists, which may themselves
CC   have use as antibacterial agents. The polynucleotides of the invention
CC   may also be used in gene therapy protocols.
XX
SQ   Sequence 969 BP; 136 A; 236 C; 317 G; 280 T; 0 other;

Query Match   22.7%; Score 15; DB 21; Length 969;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   15 GTTCGACAAAACGGT 29
         |||||||||||||||
Db 244 GTTCGACAAAACGGT 230

RESULT 42
AAF94369/c
ID   AAF94369 standard; DNA: 1104 BP.
XX
AC   AAF94369;
XX
DT   04-JUN-2001 (first entry)
XX
DE   Haemophilus influenzae essential bacterial gene SEQ ID NO:49.
XX
KW   Haemophilus influenzae; essential bacterial gene; identification;
KW   otitis media; meningitis; upper respiratory tract infection;
KW   infection; antimicrobial; ds.
XX
OS   Haemophilus influenzae.
XX
XX
PN   WO200111033-A2.
XX
PD   15-FEB-2001.
XX

PF   03-AUG-2000; 2000WO-US21176.
XX
PR   04-AUG-1999;   99US-0368382.
XX
PA   (ABBO) ABBOTT LAB.
XX
PI   Chovan LE, Hessler PE, Reich KA;
XX
DR   WPI: 2001-147511/15.
DR   P-PSDB: AAB88516.
XX
PT   Essential bacterial genes from Haemophilus influenzae and methods for
PT   identifying 'essential' genes that may be potential therapeutic targets
XX
PS   Claim 2; Page 90-91; 185pp; English.
XX
CC   AAF94345 to AAF94409 represent essential bacterial genes from
CC   Haemophilus influenzae, which encode the proteins given in AAB88492 to
CC   AAB88556. The present invention also describes methods for identifying
CC   essential bacterial genes (i.e. those essential to the survival of a
CC   bacterium) using a transposition system. The methods are used to
CC   identity essential genes from bacteria, especially H. influenzae (which
CC   causes otitis media, meningitis and upper respiratory tract infections)
CC   which may be used as targets for potential antimicrobial agents.
CC   AAF94410 to AAF94416 represent PCR primers used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 1104 BP; 347 A; 195 C; 219 G; 343 T; 0 other;

Query Match   22.7%; Score 15; DB 22; Length 1104;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   31 AAAAAGCGGTAGATT 45
         |||||||||||||||
Db 729 AAAAAGCGGTAGATT 715

RESULT 43
AAX14265
ID   AAX14265 standard; DNA: 1170 BP.
XX
AC   AAX14265;
XX
DT   31-MAR-1999 (first entry)
XX
DE   H. pylori GHPO 1564 gene.
XX
KW   GHPO protein; Helicobacter infection; gastroduodenal disease; gastritis;
KW   peptic ulcer disease; ss.
XX
OS   Helicobacter pylori.
XX
FH   Key   Location/Qualifiers
FT   CDS   97..1122
FT   /*laq- a
XX
PN   WO9843478-A1.
XX
PD   08-OCT-1998.
XX
PF   01-APR-1998; 98WO-US06371.
XX
PR   29-JUL-1997; 97US-0902615.
PR   01-APR-1997; 97US-0833457.
PR   24-JUN-1997; 97US-0881227.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (INMR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.
XX
PI   Al-Garawi A, Kleanthous H, Miller C, Oomen RP, Tomb J;
XX

```
DR   WPI; 1998-542291/46.
DR   P-PSDB; AAW98546.
XX
PT   New isolated Helicobacter polynucleotides - used to develop products
PT   for the diagnosis, prevention and treatment of Helicobacter
PT   infections and gastrointestinal diseases
XX
PS   Claim 1; Page 1053-1055; 2054pp; English.
XX
CC   This sequence represents a polynucleotide of the invention. It was
CC   isolated from Helicobacter pylori and encodes a H.pylori GHPO protein.
CC   The polypeptides can be used for preventing or treating Helicobacter
CC   infections, and gastroduodenal diseases associated with these
CC   infections, including acute, chronic, and atrophic gastritis, and peptic
CC   ulcer diseases, e.g. gastric and duodenal ulcers. They can also be used
CC   for the production of antibodies. The products can also be used for
CC   detection and diagnosis.
XX
SQ   Sequence 1170 BP; 380 A; 180 C; 272 G; 338 T; 0 other;

   Query Match       22.7%;  Score 15;  DB 19;  Length 1170;
   Host Local Similarity 100.0%;  Prod. No. 92;
   Matches  15;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

QY    29 TAAAAAAGCGGTAGA 43
          |||||||||||||||
Db   855 TAAAAAAGCGGTAGA 869

RESULT 44
AHA19747
ID   AHA19747 standard; DNA; 1457 BP.
XX
AC   AHA19747;
XX
DT   24-JAN-2002 (first entry)
XX
DE   Human nervous system related polynucleotide SEQ ID NO 12078.
XX
KW   Human; nootropic; neuroprotective; cytostatic; dermatological; viricide;
KW   immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW   antiparkinsonian; antisickling; antianaemic; antiarthritic; cancer;
KW   antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW   antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW   antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW   neurological disease; infection; nephrotropic; gene therapy; vaccine; ds.
XX
OS   Homo sapiens.
XX
PN   WO200159064-A2.
XX
PD   16-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01334.
XX
PR   41-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198124.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.

PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   24-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   08-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232497.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0231063.
PR   14-SEP-2000; 2000US-0232064.
PR   14-SEP-2000; 2000US-0233065.
PR   14-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   29-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   12-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0242221.
PR   01-NOV-2000; 2000US-0242222.
PR   08-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
```

PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 08-NOV-2000; 2000US-0249207.
PR 08-NOV-2000; 2000US-0249208.
PR 08-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249228.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250391.
PR 01-DEC-2000; 2000US-0251150.
PR 01-DEC-2000; 2000US-0251160.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 08-DEC-2000; 2000US-0254097.
PR 11-DEC-2000; 2000US-0259678.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Harash SC, Ruben SM;
XX
DR WPI; 2001-541565/60.
XX
PT Nucleic acids encoding 3224 human nervous system antigen polypeptides, useful for preventing, diagnosing and/or treating nervous system cancers and metastases -
XX
PS Disclosure; SEQ ID NO 12078; 170lpp + Sequence Listing; English.

The invention relates to novel genes (AHA11004-AHA21534) and proteins (ABB14678-ABB18001) useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The genes are isolated from a range of human tissues disclosed in the specification. The nucleic acids, proteins, antibodies and (ant)agonists are useful in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer and other cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases such as viral, bacterial, fungal and parasitic infections.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1457 BP; 383 A; 345 C; 263 G; 466 T; 0 other;

Query Match   22.7%; Score 15; DB 22; Length 1457;
Best Local Similarity   100.0%; Pred. No. 90;
Matches   15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   32 AAAAGCGGTAGANTA 46
          |||||||||||||||
Db   731 AAAAGCGGTAGANTA 745

RESULT 45
ABL02975
ID   ABL02975 standard; cDNA; 1550 BP.
XX
AC   ABL02975;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 3407.
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-1916337P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE) PE CORP NY.
XX
PI   Venter JC, Adams M, Li PWD, Myers EW;
XX
DR   WPI; 2001-6558660/75.
DR   P-PSDB; ABH58872.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions -
XX
PS   Claim 1; SEQ ID NO 3407; 21pp + Sequence Listing; English.

CC   The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA sequences (ABL16176-ABL30511) and the encoded proteins (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1550 BP; 380 A; 406 C; 408 G; 356 T; 0 other;

Query Match   22.7%; Score 15; DB 23; Length 1550;
Best Local Similarity   100.0%; Pred. No. 90;
Matches   15; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   20 ACAAAACGGTAAAAA 34
          |||||||||||||||
Db   175 ACAAAACGGTAAAAA 189

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 07:02:08 : Search time 790.981 Seconds
                                             (without alignments)
                                             1126.196 Million cell updates/sec

Title: us-09-766-113-2
Perfect score: 66
Sequence: 1 tggaatctggatccgttcga..........ccgtcctacgaattcagctg 66

Scoring table: OLIGO_NUC
                Gapop 60.0 , Gapext 60.0
Searched: 0
Word size :   0
Total number of hits satisfying chosen parameters: 27472414
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :   EST:*
                1: em_estba:*
                2: em_esthum:*
                3: em_estin:*
                4: em_estmu:*
                5: em_estov:*
                6: em_estpl:*
                7: em_estro:*
                8: em_htc:*
                9: qb_est1:*
               10: qb_est2:*
               11: qb_htc:*
               12: qb_gss:*
               13: em_gss_hum:*
               14: em_gss_inv:*
               15: em_gss_pln:*
               16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	17	25.8	456	9	AI755421	AI755421 EtEStea22
c	2	17	25.8	559	10	BE775681	BE775681 MY-05-E-1
c	3	17	25.8	568	10	BJ008710	BJ008710
c	4	17	25.8	630	10	BJ003054	BJ003054
c	5	17	25.8	723	10	BJ022467	BJ022467
c	6	17	25.8	778	12	AQ870164	AQ870164
c	7	17	25.8	927	10	BF306166	BF306166 nbeb0036H
c	8	17	25.8	211	9	HF306166	60I893015
c	9	16	24.2	305	12	BE213126	BE213126 lpbrn0156
c	10	16	24.2	427	12	TA104G02Q	TA104G02Q
c	11	16	24.2	432	12	TA104E02O	TA104E02O T. brucei
c	12	16	24.2	427	12	AL462316	AL462316 T. brucei
c	13	16	24.2	440	12	BM145840	BM145840 TCAAP1D80
c	14	16	24.2	441	9	BH170870	BH170870 SALK_0034
c	15	16	24.2	442	9	AW462196	AW462196 BP230008B
c	16	16	24.2	466	12	AW225607	AW225607 I2100586
c	17	16	24.2	478	9	AQ500181	AQ500181 V15F8 min
c	18	16	24.2	479	10	BG520697	BG520697 ps50C03.y
c	18	16	24.2	479	10	BE444883	BE444883 WHE1129 F
c	19	16	24.2	485	10	BI582395	BI582395 RH204445.5
c	20	16	24.2	486	12	B98126	B98126 F24A16TFB I
c	21	16	24.2	516	10	BI607662	BI607662 RH74766.5
c	22	16	24.2	532	10	BI584222	BI584222 RH23234.5
c	23	16	24.2	540	9	AI457082	AI457082 LD37353.5
c	24	16	24.2	563	12	AL490286	AL490286 T. brucei
c	25	16	24.2	571	10	BI612678	BI612678 RH16126.5
c	26	16	24.2	574	12	TA5H11Q	TA5H11Q T. brucei
c	27	16	24.2	580	10	BI565318	BI565318 RH63067.5
c	28	16	24.2	590	9	AI485800	AI485800 EST244121
c	29	16	24.2	600	10	BI630819	BI630819 RH60131.5
c	30	16	24.2	611	10	BI568118	BI568118 RH49495.5
c	31	16	24.2	617	10	BI237091	BI237091 RE33250.5
c	32	16	24.2	629	10	BE442560	BE442560 WHH1103 D
c	33	16	24.2	635	12	AQ950782	AQ950782 Sheared D
c	34	16	24.2	647	10	BI624045	BI624045 RH56430.5
c	35	16	24.2	649	9	AA202447	AA202447 LD02644.5
c	36	16	24.2	653	10	BI229347	BI229347 RE27113.5
c	37	16	24.2	660	9	AW690071	AW690071 NF028A065
c	38	16	24.2	660	12	AZ573452	AZ573452 31PvE10
c	39	16	24.2	668	12	BI234820	BI234820 RE30591.5
c	40	16	24.2	672	12	AZ572058	AZ572058 302PVHU3
c	41	16	24.2	673	10	BI621542	BI621542 RH52894.5
c	42	16	24.2	679	12	AZ569442	AZ569442 260PVHO2
c	43	16	24.2	681	10	BI582214	BI582214 RH20171.5
c	44	16	24.2	690	9	HH432863	HH432863 HH432863
c	45	16	24.2	694	10	HH954699	HH954699 HVSMEM001

ALIGNMENTS

RESULT 1
AI755421/c

LOCUS   AI755421   456 bp   mRNA   linear   EST 18-JAN-2000
DEFINITION EtEStea22e05.y1 Eimeria Eimeria S5-2 Sporozoite stage Eimeria tenella cDNA
            5', mRNA sequence.
ACCESSION   AI755421
VERSION   AI755421.1   GI:5149068
KEYWORDS   EST.
SOURCE   Eimeria tenella.
  ORGANISM Eimeria tenella
            Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
            Eimeria.
REFERENCE   1 (bases 1 to 456)
  AUTHORS   Liberator,P., Diaz,C., Tang,K., Marra,M., Hillier,L., Kucaba,T.,
            Martin,J., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,
            M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,
            N., Schurk,R., Ritter,E., Kohn,S., Florence,N., Shin,T., Jackson,
            Y., Cardenas,M., McCann,R., Waterston,R., Wilson,R. and Sibley,D.
  TITLE   Washu-Merck Eimeria Eimeria tenella project
  JOURNAL   Unpublished (1999)
COMMENT   Contact: David Sibley, Ph.D.
            Washu-Merck Eimeria tenella project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: estwatson.wustl.edu
            Contact:David Sibley (toxoest@porcim.wustl.edu) for further
            information relating to organism, libraries, or clone availability.
            Seq primer: -40bP from Gibco
            High quality sequence stop: 418.
FEATURES   Location/Qualifiers
     source   1..456
                     /organism="Eimeria tenella"
                     /strain="LS18"
                     /db_xref="taxon:5802"
                     /clone_lib="Eimeria S5-2 Sporozoite stage"
                     /dev_stage="Sporozoite"
                     /lab_host="Eimeria S5-2 Sporozoite stage"
                     /note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2: XhoI

; Sporozoites were obtained from in vitro sporulated and
"xcysted oocysts of E. tenella grown in chickens. cDNA
was synthesized from poly mRNA using an oligo-dT primer
containing a XhoI site. Following second strand synthesis,
EcoRI adapters were ligated to the cDNA and products were
size-selected on Sephacryl S500. cDNAs were digested with
EcoRI/XhoI and cloned into lambda Zap 11 (Stratagene).
Clones were converted to phagemids by mass excision using
ExAssist helper phage and SOLR cells (Stratagene)."

BASE COUNT   106 a   122 c   114 g   114 t
ORIGIN
                Insert sizes range from 1.2-2.9 kb."

Qy   9 GGATCCGTTCGACAAAA 25
           |||||||||||||||||
Db   44 GGATCCGTTCGACAAAA 17

Query Match   25.8%;   Score 17;   DB 9;   Length 456;
Best Local Similarity   100.0%;   Pred. No. 49;
Matches 17; Conservative   0; Mismatches   0; Indels 0; Gaps   0;

RESULT 2
BE775681
LOCUS   BE775681
DEFINITION   MY-05-E-11 PintestansMY Phytophthora infestans cDNA, mRNA sequence.
ACCESSION   BE775681
VERSION   BE775681.1 GI:10229336
KEYWORDS   EST.
SOURCE   potato late blight agent.
ORGANISM   Phytophthora infestans
              Eukaryota; stramenopiles; Oomycetes; Pythiales; Pythiaceae;
              Phytophthora.
REFERENCE   1 (bases 1 to 559)
AUTHORS   Kamoun,S., Hraber,P.T., Sobral,B.W.S., Nuss,D. and Govers,F.
TITLE   Initial assessment of gene diversity for the oomycete pathogen
              Phytophthora infestans based on expressed sequences
JOURNAL   Fungal Genet. Biol. 28 (2), 94-106 (1999)
MEDLINE   20056476
COMMENT   Contact: Govers F
              Laboratory of Phytopathology
              Wageningen University
              Binnenhaven 9, P.O.Box 8025, 6700 EE, Wageningen, The Netherlands
              Tel: 31 317 483 138
              Fax: 31 317 483 412
              Email: Francine.Govers@medew.fyto.wau.nl.
FEATURES   Location/Qualifiers
     source   1..559
                     /organism "Phytophthora infestans"
                     /strain "DDR7602, A1 mating type"
                     /db_xref "taxon:4787"
                     /clone_lib "PintestansMY"
                     /dev_stage "4-week old vegetative, non-sporulating
                     mycelium in synthetic medium"
                     /lab_host "E. coli, strain DH5-alpha"
                     /note "Vector: pSPORT1; Site_1: SalI; Site_2: NotI; Total
                     RNA was isolated from mycelium of P. infestans DDR7602
                     cultured for 4 weeks in synthetic medium. EST clones were
                     named by their position in the microtiter plate, preceded
                     by the prefix MY (for mycelial) and the successive number
                     of the microtiter plate (e.g. MY-06-A-04)."

BASE COUNT   121 a   160 c   145 g   128 t   3 others
ORIGIN

Qy   2 GAGATTCGGATCCGTTC 18
           ||||||||||||||||||
Db   225 GAGATTCGGATCCGTTC 209

Query Match   25.8%;   Score 17;   DB 10;   Length 559;
Best Local Similarity   100.0%;   Pred. No. 38;
Matches 17; Conservative   0; Mismatches   0; Indels 0; Gaps   0;

RESULT 3
BJ008710
LOCUS   BJ008710
DEFINITION   BJ008710 MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA124011 5',
              mRNA sequence.
ACCESSION   BJ008710
VERSION   BJ008710.1 GI:17356542
KEYWORDS   EST.
SOURCE   Japanese medaka.
ORGANISM   Oryzias latipes
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
              Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
              Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1 (bases 1 to 568)
AUTHORS   Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLE   Medaka EST Project in Takeda's lab
JOURNAL   Unpublished (2001)
COMMENT   Contact: Tadasu Shin-i
              Center For Genetic Resource Information
              National Institute of Genetics
              1111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tshini@genes.nig.ac.jp.
FEATURES   Location/Qualifiers
     source   1..568
                     /organism "Oryzias latipes"
                     /strain "Hd-rR"
                     /db_xref "taxon:8090"
                     /clone "MF01SSA124011"
                     /clone_lib "MF01SSA cDNA"
                     /sex "mixture of female and male"
                     /tissue_type "whole embryo"
                     /dev_stage "segmentation stage 20 - 25"

BASE COUNT   117 a   155 c   195 g   101 t
ORIGIN

Qy   4 AGATCTGGATCCGTTCG 19
           ||||||||||||||||||
Db   425 AGATCTGGATCCGTTCG 441

Query Match   25.8%;   Score 17;   DB 10;   Length 568;
Best Local Similarity   100.0%;   Pred. No. 38;
Matches 17; Conservative   0; Mismatches   0; Indels 0; Gaps   0;

RESULT 4
BJ003054
LOCUS   BJ003054
DEFINITION   BJ003054 MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA042E04 5',
              mRNA sequence.
ACCESSION   BJ003054
VERSION   BJ003054.1 GI:17365174
KEYWORDS   EST.
SOURCE   Japanese medaka.
ORGANISM   Oryzias latipes
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
              Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
              Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1 (bases 1 to 630)
AUTHORS   Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
TITLE   Medaka EST Project in Takeda's lab
JOURNAL   Unpublished (2001)
COMMENT   Contact: Tadasu Shin-i
              Center For Genetic Resource Information
              National Institute of Genetics
              1111 Yata, Mishima, Shizuoka 411-8540, Japan
              Tel: 81-559-81-6856
              Fax: 81-559-81-6855
              Email: tshini@genes.nig.ac.jp.
FEATURES

source   1..630
                /organism="Oryzias latipes"
                /strain="Hd-rR"
                /db_xref="taxon:8090"
                /clone="MF01SSA042E04"
                /clone_lib="MF01SSA cDNA"
                /sex="mixture of female and male"
                /tissue_type="whole embryo"
                /dev_stage="segmentation stage 20 - 25"

BASE COUNT   144 a   161 c   214 g   111 t
ORIGIN

Qy   3 AGAATCTGGATCCGTTCG 19
              |||||||||||||||||||
Db   575 AGAATCTGGATCCGTTCC 591

RESULT 5
LOCUS   HJ022467
DEFINITION BJ022467 MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA123D11 3',
            mRNA sequence.
ACCESSION   BJ022467
VERSION   BJ022467.1   GI:17373869
KEYWORDS   EST.
SOURCE   Japanese medaka.
  ORGANISM Oryzias latipes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1 (bases 1 to 723)
  AUTHORS   Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
  TITLE   Medaka EST Project in Takeda's lab
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Tadasu Shin-i
            National Institute of Genetics
            Center For Genetic Resource Information
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshin-i@genes.nig.ac.jp.
FEATURES   Location/Qualifiers
     source   1..723
                     /organism="Oryzias latipes"
                     /strain="Hd-rR"
                     /db_xref="taxon:8090"
                     /clone="MF01SSA123D11"
                     /clone_lib="MF01SSA cDNA"
                     /sex="mixture of female and male"
                     /tissue_type="whole embryo"
                     /dev_stage="segmentation stage 20 - 25"
BASE COUNT   152 a   229 c   196 g   145 t   1 others
ORIGIN

Qy   3 AGAATCTGGATCCGTTCG 19
              |||||||||||||||||||
Db   537 AGAATCTGGATCCGTTCG 521

RESULT 6
LOCUS   AQ870164/c
DEFINITION AQ870164 CUGI Rice BAC Library (EcoRI) Oryza sativa genomic
            clone nbeb0036H20r, DNA sequence.

LOCUS   BJ022467   723 bp   mRNA   linear   EST 05-DEC-2001

Query Match   25.8%; Score 17; DB 10; Length 630;
Best Local Similarity   100.0%; Pred. No. 37;
Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

DEFINITION MF01SSA cDNA Oryzias latipes cDNA clone MF01SSA123D11 3',
            mRNA sequence.

Query Match   25.8%; Score 17; DB 10; Length 723;
Best Local Similarity   100.0%; Pred. No. 36;
Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

LOCUS   AQ870164   778 bp   DNA   linear   GSS 03-NOV-1999

ACCESSION   AQ870164
VERSION   AQ870164.1   GI:6220615
KEYWORDS   GSS.
SOURCE   Oryza sativa
  ORGANISM Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE   1 (bases 1 to 778)
  AUTHORS   Wing,R.A. and Dean,R.A.
  TITLE   A BAC End Sequencing Framework to Sequence the Rice Genome
  JOURNAL   Unpublished (1998)
COMMENT   Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7298
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq_primer: GGAAACAGGTATGACCATG
            Class: BAC ends
            /lab_host="E. coli DH10B"
            /tissue_type="leaf"
            /note="Vector: pHACIndigo; Site_1: EcoRI; Site_2: EcoRI;
            Rice is the most important food crop in the world. Half of
            the world population, especially those inhabiting highly
            populated areas of the humid tropics and subtropics, rely
            on rice as their primary source of carbohydrate.
            Monocotyledonous rice is a diploid plant (2n=24) with a
            haploid genome equivalent of 431 Mbp (Arumuganathan and
            Earle, 1991). The relatively small genome of rice, three
            times larger than that of Arabidopsis, makes it suitable
            for genomic studies. In order to facilitate positional
            cloning, physical mapping and genome sequencing of rice,
            we have constructed a BAC library from Oryza sativa,
            Nipponbare variety using EcoRI as the cloning enzyme. The
            library contains 55,296 clones with an average insert size
            of 121 kb providing approximately 15 haploid genome
            equivalents. The deep coverage allows the isolation a
            particular sequence with a probability of 99.9 %. Three
            high density filters, each containing 18,432 clones
            (doubly spotted), represent the whole library for colony
            screening and can be requested from the Clemson University
            BAC/EST Resource Center (www.genome.clemson.edu).
FEATURES   Location/Qualifiers
     source   1..778
                     /organism="Oryza sativa"
                     /strain="Japonica"
                     /cultivar="Nipponbare"
                     /db_xref="taxon:4530"
                     /clone="nbeb0036H20r"
                     /clone_lib="CUGI Rice BAC Library (EcoRI)"
            High quality sequence start: 29
            High quality sequence stop: 522.
BASE COUNT   265 a   130 c   138 g   243 t   2 others
ORIGIN

Qy   19 GACAAAACGGTAAAAAA 35
              |||||||||||||||||
Db   346 GACAAAACGGTAAAAAA 330

RESULT 7
LOCUS   BF306166
DEFINITION 601893015F1 NIH_MGC_17 Homo sapiens cDNA clone IMAGE:4138922 5',
            mRNA sequence.
ACCESSION   BF306166
VERSION   BF306166.1   GI:11253325

LOCUS   BF306166   927 bp   mRNA   linear   EST 21-NOV-2000

Query Match   25.8%; Score 17; DB 12; Length 778;
Best Local Similarity   100.0%; Pred. No. 36;
Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

KEYWORDS   EST.
SOURCE   human.
ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 927)
AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
             Plate: LLCM1048 row: m column: 03.
FEATURES   Location/Qualifiers
     source   1..927
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4149922"
                     /clone_lib="NIH_MGC_17"
                     /tissue_type="rhabdomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: muscle; Vector: pOTB7; Site_1: EcoRI;
                     Site_2: XhoI; cDNA made by oligo-dT priming;
                     directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G). Size-selected >500bp
                     for average insert size 1.8kb. Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT   239 a   221 c   271 g   195 t   1 others
ORIGIN

Qy   21 AAACCGTAAAAAGCGG 39
            ||||||||||||||||
Db   814 AAACCGTAAAAAGCGG 829

RESULT H
BE211126
BE211126
LOCUS   BE211126   211 bp   mRNA   linear   EST 30-JUN-2000
DEFINITION   IpBra01569 Brain cDNA library Ictalurus punctatus cDNA 5', mRNA
             sequence.
ACCESSION   BE211126
VERSION   BE211126.1 GI:8844872
KEYWORDS   EST.
SOURCE   channel catfish.
ORGANISM   Ictalurus punctatus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
             Ictaluridae; Ictalurus.
REFERENCE   1 (bases 1 to 211)
AUTHORS   Ju,Z., Karsi,A., Kocabas,A., Patterson,A., Li,P., Cao,D., Dunham,R.
             and Liu,Z.
TITLE   Transcriptome analysis of channel catfish: I. genes and expression
             profiles from the brain
JOURNAL   Unpublished (2000)
COMMENT   Contact: Liu, Z.J.
             Fish Molecular Genetics and Biotechnology
             Auburn University
             203 Swingle Hall, Department of Fisheries, Auburn, AL 36849, USA
             Tel: 334 844 4054
             Fax: 334 844 9208
             Email: zliu@acesag.auburn.edu
             Seq primer: M13 Reverse.

Query Match   25.4%; Score 17; DB 10; Length 927;
Best Local Similarity   100.0%; Pred. No. 35;
Matches   17; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

FEATURES   Location/Qualifiers
     source   1..211
                     /organism="Ictalurus punctatus"
                     /db_xref="taxon:7998"
                     /clone_lib="Brain cDNA library"
                     /note="Organ: Brain; Vector: pSport1; Site_1: NotI;
                     Site_2: SalI"
BASE COUNT   53 a   33 c   49 g   76 t
ORIGIN

Qy   20 ACAAAAACGGTAAAAAA 35
            ||||||||||||||||
Db   192 ACAAAAACGGTAAAAAA 207

RESULT 9
TA104G02Q/c
LOCUS   TA104G02Q   305 bp   DNA   linear   GSS 14-DEC-2000
DEFINITION   T. brucei sheared genomic DNA clone 104g02, reverse sequence,
             genomic survey sequence.
ACCESSION   AL462316
VERSION   AL462316.1 GI:11831561
KEYWORDS   GSS.
SOURCE   Trypanosoma brucei
ORGANISM   Trypanosoma brucei
             Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
             Trypanosoma.
REFERENCE   1 (bases 1 to 305)
AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
             Chillingworth,C., Ormond,D., Harris,B., El Sayed,N., Hou,L.,
             Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE   Direct Submission
JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
             project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
             Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
             nh1@sanger.ac.uk
COMMENT   Constructed at the Institute for Genomic Research (TIGR),
             Rockville, MD. Genomic DNA isolated from a cloned population of
             Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
             to give a tight size distribution (
             4 kb). The v + i method used for the library construction is
             described in detail in Smith, H. and Venter, J.C. (Making small
             insert libraries for whole genome shotgun sequencing projects. In
             Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
             Barrell, Oxford University Press, 1999).
             Email: nelsayed@tigr.org
             Details of T. brucei sequencing at the Sanger Centre are available
             at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES   Location/Qualifiers
     source   1..305
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="104g02"
BASE COUNT   75 a   69 c   54 g   107 t
ORIGIN

Qy   20 ACAAAACGGTAAAAAAA 35
            ||||||||||||||||
Db   116 ACAAAACGGTAAAAAAA 101

RESULT 10
TA104E02Q/c
LOCUS   TA104E02Q   427 bp   DNA   linear   GSS 13-DEC-2000

Query Match   24.2%; Score 16; DB 12; Length 305;
Best Local Similarity   100.0%; Pred. No. 1.5e+02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Query Match   24.2%; Score 16; DB 9; Length 211;
Best Local Similarity   100.0%; Pred. No. 1.6e+02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

DEFINITION   T. brucei sheared genomic DNA clone 104c02, reverse sequence,
              genomic survey sequence.
ACCESSION   AL462307
VERSION   AL462307.1 GI:11831552
KEYWORDS   GSS.
SOURCE   Trypanosoma brucei.
  ORGANISM   Trypanosoma brucei
              Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
              Trypanosoma.
REFERENCE   1 (bases 1 to 427)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
              Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
              Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE   Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
              project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
              Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
              nhl@sanger.ac.uk
COMMENT   Constructed at the Institute for Genomic Research (TIGR),
              Rockville, MD. Genomic DNA isolated from a cloned population of
              Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
              to give a tight size distribution (
              4 kb). The v + i method used for the library construction is
              described in detail in Smith, H. and Venter, J.C. (Making small
              insert libraries for whole genome shotgun sequencing projects. In
              Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
              Barrell, Oxford University Press, 1999).
              Email: nelsayed@tigr.org
              Details of T. brucei sequencing at the Sanger Centre are available
              at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES   Location/Qualifiers
     source   1..427
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="104e02"
BASE COUNT   129 a   90 c   84 g   123 t   1 others
ORIGIN

Qy   20 ACAAAACGGTAAAAAA 35
              ||||||||||||||||
Db   125 ACAAAACGGTAAAAAA 110

Query Match   24.2%; Score 16; DB 12; Length 427;
Best Local Similarity 100.0%;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 11
LOCUS   BM145840
BM145840/c
DEFINITION TCAAP1D8023 Pediatric acute myelogenous leukemia cell (FAB
            M1) Baylor-HGSC project=TCAA Homo sapiens cDNA clone TCAAP8023, mRNA
            sequence.
                     432 bp   mRNA   linear   EST 30-NOV-2001
ACCESSION   BM145840
VERSION   BM145840.1 GI:17164228
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 432)
  AUTHORS   Wei,Y., Tsang,Y.T.M., Mei,G., Ku,J.M., Ali-Osman,F.R.Jr.,
            Guaratne,P.H., Muzny,D., Bouck,J., Gibbs,R.A. and Margolin,J.F.
  TITLE   Pediatric Leukemia cDNA Sequencing Project (2001)
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Dr. Judith F. Margolin
            Texas Children's Cancer Center and Human Genome Sequencing Center
            at Baylor College of Medicine
            1102 Bates, MC3-3320 Houston, TX 77030, USA
            Tel: 832-824-4536
            Fax: 832-825-4038
            Email: clones@txccc.org

FEATURES   Location/Qualifiers
     source   1..432
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="TCAAP8023"
                     /clone_lib="Pediatric acute myelogenous leukemia cell (FAB
                     M1) Baylor-HGSC project=TCAA"
                     /sex="male"
                     /tissue_type="leukopheresis"
                     /cell_type="myeloid cell"
                     /dev_stage="pediatric 6 years"
                     /lab_host="DH10B"
                     /note="Vector: lambda pSB; Site_1: BamHI; Site_2: EcoRI;
                     First strand cDNA was primed with an anchored
                     XhoI oligo(dT) primer [5'GGAGACTCGAGCGGCCGCGCCGAGGAGGAG(T)VN
                     3'; V=A,C,G; N=A,C,G,T] and then dC tailed. Second strand
                     was primed with a BamHI-dC primer
                     [5'AGAGAGCTCGGATCCGGGCCGCCCAATAATAATAAT(C) 3'].
                     Double-stranded cDNA was then digested with BamHI and XhoI
                     and directionally cloned into the BamHI and SalI sites of
                     lambda pSB vector. Library went through one round of
                     normalization. Library was constructed by Wei Yu at RIKEN
                     of Japan (Carninci P, Westover A, Nishiyama Y, Ohsumi T,
                     Itoh M, Nagaoka S, Sasakin K, Okazaki Y, Muramatsu M,
                     Schneider C, Hayashizaki Y. High efficiency selection of
                     full-length cDNA by improved biotinylated cap trapper...
                     DNA Res 4:1, 61-6, Feb 28, 1997)"
Seq primer: M13 primer

Query Match   24.2%; Score 16; DB 10; Length 432;
Best Local Similarity 100.0%;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;
BASE COUNT   97 a   74 c   88 g   173 t
ORIGIN

Qy   22 AAAACGGTAAAAAGC 37
              |||||||||||||||
Db   187 AAAACGGTAAAAAGC 172

RESULT 12
LOCUS   BH170870
BH170870/c
DEFINITION SALK_003469 Arabidopsis thaliana TDNA insertion lines Arabidopsis
            thaliana genomic clone SALK_003469, DNA sequence.
                     440 bp   DNA   linear   GSS 03-OCT-2001
ACCESSION   BH170870
VERSION   BH170870.1 GI:15906625
KEYWORDS   GSS.
SOURCE   thale cress.
  ORGANISM Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 440)
  AUTHORS   Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadrinab
            ,C., Heller,C., Kim,C.J., Koesema,F., Meyers,M.C., Parker
            ,H., Prednis,L., Shinn,P., Stevenson,D.K., Zimmerman,J. and Ecker
            ,J.R.
  TITLE   A Sequence-Indexed Library of Insertion Mutations in the
            Arabidopsis Genome
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Joseph R. Ecker
            Salk Institute Genomic Analysis Laboratory (SIGNAL)
            The Salk Institute for Biological Studies
            10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
            Tel: 858 453 4100 x1752
            Fax: 858 558 6379
            Email: ecker@salk.edu
            This is single pass sequence recovered from the left border of
            TDNA.
            Class: TDNA tagged.
FEATURES   Location/Qualifiers
     source   1..440

```
          /organism="Arabidopsis thaliana"
          /strain="columbia 0"
          /db_xref="taxon:3702"
          /clone="SALK_0034469"
          /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
          /note="PCR was performed on Arabidopsis thaliana lines
          each of which contains one or more T-DNA insertion
          elements. The resultant fragment for each line was
          directly sequenced to determine the genomic sequence at
          the site of insertion. Details of the protocols used can
          be found at http://signal.salk.edu/tdna_protocols.html"

Qy    20 ACAAAACGGTAAAAAA 35
          ||||||||||||||||
Db   247 ACAAAACGGTAAAAAA 222

BASE COUNT     122 a    112 c    79 g    127 t
ORIGIN

Query Match       24.2%;   Score 16;   DB 12;   Length 440;
Best Local Similarity 100.0%;     Pred. No. 1.4e+02;
Matches 16;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

RESULT 14
AW46219b/c
LOCUS       AW462196      441 bp    mRNA    linear   EST 24-FEB-2000
DEFINITION  HP230008H10G12 Soares normalized bovine placenta Bos taurus cDNA
            clone HP230008H10G12 5', mRNA sequence.
ACCESSION   AW462196
VERSION     AW462196.1  GI:7012364
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 441)
  AUTHORS   Lewin,H.A., Soares,M.B., Rebeiz,M., Pardinas,J., Liu,L. and Larson,J.H.
  TITLE     Bovine ESTs
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Lewin, H. A.
            W. M. Keck Center for Comparative and Functional Genomics
            University of Illinois at Urbana-Champaign
            340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL
            61801, USA
            Tel: 217 333 5998
            Fax: 217 244 5617
            Email: h-lewin@uiuc.edu
            Funding for cattle EST sequencing was provided by the USDA National
            Research Initiative, Animal Genome Resource Grant AG 99-35205-8534
            to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
            from Washington University Genome Center. Vector Trimming: PHRAP
            suite. Cross_match from Washington University Genome Center PHRAP
            suite. Sequences submitted are vector free and at least 200 bp in length.
            PCR primers
            FORWARD: TAATACGACTCACTATAGGG
            BACKWARD: ATTAACCCTCACTAAAG
            Insert Length: 441    Std Error: 0.00
            Plate: HP230008H10    row: G  column: 12
            Seq primer: ACGGATAACAATTTCACACAGGA
            High quality sequence stop: 441.
FEATURES             Location/Qualifiers
     source          1..441
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone="HP230008H10G12"
                     /clone_lib="Soares normalized bovine placenta"
                     /sex="female"
                     /lab_host="DH10B"
                     /note="Organ: placenta; Vector: pT7T3Pac; Site_1: EcoRI;
                     Site_2: NotI; The cDNA library was contributed by the
                     Soares laboratory and it was constructed and normalized
```

```
            as described by Bonaldo, M.F., Lennon, G. and Soares,
            M.B. (1996), Genome Research 6(9): 791-806.

BASE COUNT     131 a    77 c    94 g    139 t
ORIGIN

Query Match       24.2%;   Score 16;   DB 9;   Length 441;
Best Local Similarity 100.0%;     Pred. No. 1.4e+02;
Matches 16;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy    50 TCCTACGAATTCAGCT 65
          ||||||||||||||||
Db    39 TCCTACGAATTCAGCT 24

RESULT 14
AW225607/c
LOCUS       AW225607      442 bp    mRNA    linear   EST 05-SEP-2000
DEFINITION  T2100580 KV0 Medicago truncatula cDNA clone pKV0-2J4, mRNA
            sequence.
ACCESSION   AW225607
VERSION     AW225607.1  GI:6554903
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 442)
  AUTHORS   VandenBosch,K., Hur,J., Beremand,P., Peng,H. and Ellis,L.
  TITLE     ESTs from uninoculated roots of Medicago truncatula
  JOURNAL   Unpublished (1999)
COMMENT     Contact: VandenBosch K
            Department of Biology
            Texas A&M University
            College Station, TX 77843-3258, USA
            Tel: 409 845 7707
            Fax: 409 845 2891
            Email: Kate@mail.bio.tamu.edu
            Other name: 38-KV0-2J4; date: 12/1/99; Submitted to the database of
            Expressed Sequence Tags (dbEST) on 12/10/99; More information is
            available at 'http://chrysie.tamu.edu/medicago'.
            Seq primer: SKmod (CTA gAA CTA gTg gAT CC).
FEATURES             Location/Qualifiers
     source          1..442
                     /organism="Medicago truncatula"
                     /cultivar="genotype A17"
                     /db_xref="taxon:3880"
                     /clone="pKV0-2J4"
                     /clone_lib="KV0"
                     /tissue_type="Seedling roots"
                     /dev_stage="Immediately prior to inoculation with
                     Sinorhizobium meliloti (0 hour)"
                     /lab_host="E.coli strain XL01R"
                     /note="Vector: pBluescript SK -; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA enriched RNA. The cDNA
                     was directionally ligated into the Uni-Zap XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts. Plasmids containing cDNA inserts were excised
                     from the recombinant lambda Zap phage using Ex-assist
                     helper phage and propagated in XL01R cells."

BASE COUNT     133 a    124 c    50 g    135 t
ORIGIN

Query Match       24.2%;   Score 16;   DB 9;   Length 442;
Best Local Similarity 100.0%;     Pred. No. 1.4e+02;
Matches 16;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

Qy     2 GAGATCTGGATCCGTT 17
          ||||||||||||||||
Db   395 GAGATCTGGATCCGTT 380
```

RESULT 15
LOCUS   AQ500181
DEFINITION V15F8 mTn-3xHA/lacZ Insertion Library Saccharomyces cerevisiae genomic 5', DNA sequence.
ACCESSION   AQ500181
VERSION   AQ500181.1 GI:4705003
KEYWORDS   GSS.
SOURCE   baker's yeast.
ORGANISM   Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomyces.
REFERENCE   1 (bases 1 to 466)
AUTHORS   Ross-Macdonald,P., Coelho,P.S.R., Agarwal,S., Kumar,A.,
            desEtages,S.A., Cheung,K.-H., Sheehan,A., Symoniatis,D., Jansen,R.,
            Umansky,L., Heidtman,M., Nelson,K., Iwasaki,H., Kanada,D., Lugo,R.,
            Hager,K., Miller,P., Roeder,G.S. and Snyder,M.
TITLE   Large-Scale Analysis of the Yeast Genome by Transposon Tagging and
            Gene Disruption
JOURNAL   Unpublished (1999)
COMMENT   Contact: Kumar A
            Michael Snyder, Dept. of Mol. Cell. and Dev. Biology
            Yale University
            P.O. Box 208103, New Haven, CT 06520-8103, USA
            Tel: 203 432 9949
            Fax: 203 432 6161
            Email: anuj.kumar@yale.edu
            Seq primer: GGCCTTCGTTTGTTGGAAGTAC
            te of mTn-3xHA/lacZ insertion.
FEATURES   Location/Qualifiers
     source   1..466
                     /organism="Saccharomyces cerevisiae"
                     /db_xref="taxon:4932"
                     /clone_lib="mTn-3xHA/lacZ Insertion Library"
                     /lab_host="E. coli"
                     /note="Vector: pHSS6-Sal; A yeast genomic DNA library
                     (lacking mitochondrial DNA) was prepared in pHSS6-Sal;
                     genomic DNA was size-fractionated (DNA of roughly 2-3 kb
                     in length) prior to cloning. This library was
                     subsequently mutagenized with a mTn-3xHA/lacZ
                     minitransposon containing lacZ, URA3, and tet resistance."
BASE COUNT   177 a   118 c   69 g   101 t   1 others
ORIGIN

Qy   20 ACAAAACGGTAAAAA 35
             |||||||||||||||
Db   280 ACAAAACGGTAAAAA 295

Query Match   24.2%; Score 16; DB 12; Length 466;
Best Local Similarity   100.0%; Pred. No. 1.4e-02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 16
LOCUS   AA696283
DEFINITION GM05449_5prime GM Drosophila melanogaster ovary BlueScript
            Drosophila melanogaster cDNA clone GM05449 5prime, mRNA sequence.
ACCESSION   AA696283
VERSION   AA696283.1 GI:2698903
KEYWORDS   EST.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 478)
AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
TITLE   BDGP/HHMI Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 54 row: E column: 1
            High quality sequence stop: 477.
FEATURES   Location/Qualifiers
     source   1..478
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="GM05449"
                     /clone_lib="GM Drosophila melanogaster ovary BlueScript"
                     /sex="female"
                     /dev_stage="newly eclosed females; germarium-stage 6"
                     /lab_host="SOLR"
                     /note="Organ: ovary; Vector: BlueScript SK; Site_1: EcoRI;
                     Site_2: XhoI; Constructed using Stratagene ZAP-cDNA
                     Synthesis kit; Oligo dT-primed and directionally cloned at
                     EcoRI and XhoI in BlueScript SK(+/-)"
BASE COUNT   136 a   122 c   101 g   119 t
ORIGIN

Qy   19 CACAAAACGGTAAAAA 34
             ||||||||||||||||
Db   297 CACAAAACGGTAAAAA 312

Query Match   24.2%; Score 16; DB 9; Length 478;
Best Local Similarity   100.0%; Pred. No. 1.4e-02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 17
LOCUS   BG520697
DEFINITION ps50c03.y1 Trichinella spiralis ML CMVsport jasmer Trichinella
            spiralis cDNA 5', mRNA sequence.
ACCESSION   BG520697
VERSION   BG520697.1 GI:13536124
KEYWORDS   EST.
SOURCE   Trichinella spiralis
ORGANISM   Trichinella spiralis
            Eukaryota; Metazoa; Nematoda; Enoplea; Trichocephalida;
            Trichinellidae; Trichinella.
REFERENCE   1 (bases 1 to 479)
AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J., Wylie,T.,
            Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,Y.,
            Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Isgaragishvili,R.,
            Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe,
            M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,
            Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and
            Wilson,R.
TITLE   The Washington Univ. Nematode EST Project, 1999
JOURNAL   Unpublished (1999)
COMMENT   Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Dr. Doug Jasmer
            (djasmer@vetmed.wsu.edu) at Washington State University, Dept. of
            Veterinary Microbiology and Pathology DNA Sequencing by: Washington
            University Genome Sequencing Center St. Louis.
            Seq primer: Sp6.
FEATURES   Location/Qualifiers
     source   1..479
                     /organism="Trichinella spiralis"
                     /db_xref="taxon:6334"
                     /clone_lib="Trichinella spiralis ML CMVsport jasmer"
                     /dev_stage="muscle stage larvae"
                     /lab_host="DH10B"

RESULT 18

```
Qy          40 AAAAAAGCGTTAGATT 45
               ||||||||||||||||
Db         400 AAAAAAGCGTTAGATT 415
```

```
BASE COUNT    158 a        80 c       109 g      142 t
ORIGIN
```

```
Query Match            24.2%;   Score 16;   DB 10;   Length 479;
Best Local Similarity 100.0%;   Pred. No. 1.4e+02;
Matches 16;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
```

```
Qy          40 AAAAAAGCGTTAGATT 45
               ||||||||||||||||
Db         400 AAAAAAGCGTTAGATT 415
```

ACCESSION   BE444884
VERSION   BE444884.1 GI:9444435
KEYWORDS   EST.
SOURCE   bread wheat.
  ORGANISM Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 479)
  AUTHORS   Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han
            ,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Nguyen,H.T.,
            Rausch,C.J., Seaton,G.J., Tong,J.C. and Zhang,D.
  TITLE   The structure and function of the expressed portion of the wheat
            genomes - Normalized root cDNA library
  JOURNAL   Unpublished (2000)
COMMENT   Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595577;
            Fax: 5105595818
            Email: oanders@pw.usda.gov
            Seq primer: Stratagene SK primer.
FEATURES   Location/Qualifiers
     source   1..479
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE1129_F06_K11"
                     /clone_lib="Wheat etiolated seedling root normalized cDNA
                     library"
                     /tissue_type="Root"
                     /dev_stage="Five day old etiolated seedling"
                     /lab_host="E. coli DH10B"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid
                     pBluescript SK; Site_1: EcoRI; Site_2: XhoI; Seeds were
                     surface-sterilized, germinated and grown aseptically in
                     the dark at room temperature on filter paper with water,
                     crystalin and cefotaxime in covered crystallization
                     dishes. Roots were harvested. The tissue, total RNA, and
```

LOCUS       BE444884               479 bp    mRNA    linear   EST 25-JUL-2000
DEFINITION  WHE1129_F06_K11ZS Wheat etiolated seedling root normalized cDNA
            library Triticum aestivum cDNA clone WHE1129_F06_K11, mRNA
            sequence.

/note="Vector: pCMVsport-7.neo; Site_1: NotI; Site_2: SalI
; The library was constructed using mRNA isolated from
total RNA with oligo-dT cellulose. Total RNA was generated
from muscle larvae that were isolated from infected rats.
Larvae were liberated by pepsin/HCl digestion, incubated
with 18 SDS, treated with RNase and DNase to eliminate
host nucleic acid contamination, and purified on a Percoll
gradient. The T. spiralis isolate was obtained from
Dickson Despommier (Columbia University). The library was
provided by Dr. Doug Jasmer (djasmer@vetmed.wsu.edu) and
colleagues at Washington State University. DNA Sequencing
by: Washington University Genome Sequencing Center St.
Louis."

---

RESULT 19

```
Qy           7 CTGGATCCGTTCGACA 22
               ||||||||||||||||
Db         225 CTGGATCCGTTCGACA 240
```

```
BASE COUNT    119 a       114 c       126 g      120 t
ORIGIN
```

```
Query Match            24.2%;   Score 16;   DB 10;   Length 479;
Best Local Similarity 100.0%;   Pred. No. 1.4e+02;
Matches 16;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
```

ACCESSION   BI582395
VERSION   BI582395.1 GI:15473817
KEYWORDS   EST.
SOURCE   fruit fly.
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 485)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE   BDGP/HHMI RH Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST; est@fruitfly.berkeley.edu
            Plate: RH.204 row: D column: 9
            High quality sequence stop: 413.
FEATURES   Location/Qualifiers
     source   1..485
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH20445"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head; Vector: pFlcI; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."

BASE COUNT   131 a   121 c   106 g   127 t
ORIGIN
```

LOCUS   BI582395   485 bp   mRNA   linear   EST 06-SEP-2001
DEFINITION RH20445.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH20445.5 similar to yiu:
            FBan0002913 GO:[integral plasma membrane protein (GO:0005887);
            hydrogen/oligopeptide symporter (GO:0005427); hydrogen/oligopeptide
            symporter (GO:0005427); transporter (GO:0005215); integral plasma
            membrane protein (GO:0005887)]; bny. mRNA sequence.

poly(A) RNA were prepared, a cDNA library was made in the
TJ Close lab (Choi, Close, Fenton) at the University of
California, Riverside. The cDNA clones were in vivo
excised to give pBluescript phagemids before
normalization was carried out. The mass excision of
phagemid library and normalization were done in HT Nguyen
lab by D. Zhang at Texas Tech University. Normalization
protocol used was that of Soares. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."

RESULT 20

Query Match   24.2%; Score 16; DB 10; Length 485;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    19 GACAAAACGGTAAAAA 34
         ||||||||||||||||
Db   340 GACAAAACGGTAAAAA 355
```

LOCUS   B98126   486 bp   DNA   linear   GSS 31-MAR-1998
DEFINITION F24A16TFB IGF Arabidopsis thaliana genomic clone F24A16, DNA
            sequence.
ACCESSION   B98126
VERSION   B98126.1   GI:3000205
KEYWORDS   GSS.
SOURCE   thale cress.
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 486)
AUTHORS   Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K., Golden,K.
            , Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter
            ,J.C.
TITLE   A BAC End Sequence Database for Identifying Minimal Overlaps in
            Arabidopsis Genomic Sequencing. Update 3
JOURNAL   Unpublished (1997)
COMMENT   Contact: Steve Rounsley
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: rounsley@tigr.org
            Seq primer: M13-21
            Class: BAC ends
            High quality sequence stop: 486.
FEATURES   Location/Qualifiers
     source   1..486
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="F24A16"
                     /clone_lib="IGF"
                     /sex="hermaphrodite"
                     /note="Vector: BeloBACII; Site_1: EcoRI; Site_2: EcoRI;
                     Produced by Thomas Altmann"
BASE COUNT   149 a   88 c   90 g   159 t
ORIGIN

RESULT 21

Query Match   24.2%; Score 16; DB 12; Length 486;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    18 CGACAAAACGGTAAAA 33
         ||||||||||||||||
Db   380 CGACAAAACGGTAAAA 395
```

LOCUS   BI607662   516 bp   mRNA   linear   EST 07-SEP-2001
DEFINITION RH74766.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH74766 5 similar to yin:
            FBan0002913 GO:|integral plasma membrane protein (GO:0005887);
            hydrogen/oligopeptide symporter (GO:0005427); transporter
            (GO:0005215); integral plasma
            membrane protein (GO:0005887); hy, mRNA sequence.
ACCESSION   BI607662
VERSION   BI607662.1   GI:15501487
KEYWORDS   EST.

SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 516)
AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,R., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
TITLE   BDGP/HHMI RH Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: RH.747 row: F column: 6
            High quality sequence stop: 414.
FEATURES   Location/Qualifiers
     source   1..516
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH74766"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organi:head; Vector: pFlrl; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT   142 a   130 c   112 g   132 t
ORIGIN

RESULT 22

Query Match   24.2%; Score 16; DB 10; Length 516;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    19 GACAAAACGGTAAAAA 34
         ||||||||||||||||
Db   340 GACAAAACGGTAAAAA 355
```

LOCUS   BI584222   532 bp   mRNA   linear   EST 06-SEP-2001
DEFINITION RH23234.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH23234 5 similar to yin:
            FBan0002913 GO:|integral plasma membrane protein (GO:0005887);
            hydrogen/oligopeptide symporter (GO:0005427); transporter
            (GO:0005215); hydrogen/oligopeptide
            membrane protein (GO:0005887); hy, mRNA sequence.
ACCESSION   BI584222
VERSION   BI584222.1   GI:15475644
KEYWORDS   EST.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 532)
AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Herman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S., and Rubin
            ,G.M.
TITLE   BDGP/HHMI RH Drosophila EST Project

JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            one Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510-486-6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: RH_232 row: C column: 10
            High quality sequence stop: 482.
FEATURES   Location/Qualifiers
     source   1..542
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH23244"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head; Vector: pFlc; Site_1: XhoI; Site_2:
                     BamHI. Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     cre recombinase. Plasmid cDNA library."

BASE COUNT   147 a   135 c   115 g   114 t   1 others
ORIGIN

RESULT 23
LOCUS   AI457082
Qy   19 GACAAAAACGGTAAAAA 34
        ||||||||||||||||||
Db 140 GACAAAACGGTAAAAA 355

Query Match   24.2%; Score 16; DB 10; Length 532;
Best Local Similarity   100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

LOCUS   AI457082   540 bp   mRNA   linear   EST 19-APR-2001
DEFINITION LD37353.5prime LD Drosophila melanogaster embryo poT2 Drosophila
            melanogaster cDNA clone LD37353 5prime, mRNA sequence.
ACCESSION   AI457082
VERSION   AI457082.1   GI:4285451
KEYWORDS   EST.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 540)
AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
TITLE   HHMI/BDGP Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            one Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510-486-6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 374 row: E column: 5
            High quality sequence stop: 499.
FEATURES   Location/Qualifiers
     source   1..540
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="LD37353"
                     /clone_lib="LD Drosophila melanogaster embryo poT2"
                     /sex="male and female"
                     /dev_stage="0 to 24 hours mixed stage embryonic"
                     /lab_host="XL1 Blue"
                     /note="Organ: embryo; Vector: poT2; Site_1: EcoRI; Site_2:
                     XhoI. Sized fractionated cDNAs were directly ligated into

BASE COUNT   148 a   140 c   115 g   137 t
ORIGIN

RESULT 24
LOCUS   TA313B04P
Qy   19 GACAAAAACGGTAAAAA 34
        ||||||||||||||||||
Db 320 GACAAAACGGTAAAAA 335

Query Match   24.2%; Score 16; DB 9; Length 540;
Best Local Similarity   100.0%; Pred. No. 1.4e+02;
Matches 16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

LOCUS   TA313B04P   563 bp   DNA   linear   GSS 14-DEC-2000
DEFINITION T. brucei sheared genomic DNA clone 313b04, forward sequence,
            genomic survey sequence.
ACCESSION   AL490286
VERSION   AL490286.1   GI:11866376
KEYWORDS   GSS.
SOURCE   Trypanosoma brucei.
ORGANISM   Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1 (bases 1 to 563)
AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
TITLE   Direct Submission
JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT   Constructed at the Institute for Genomic Research (TIGR),
            Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb). The v-i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            Insert libraries for whole genome shotgun sequencing projects. In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nc1@sayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES   Location/Qualifiers
     source   1..563
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="313b04"

BASE COUNT   238 a   75 c   109 g   139 t   2 others
ORIGIN

RESULT 25
LOCUS   H1612678
Qy   20 ACAAAAACGGTAAAAAA 35
        ||||||||||||||||||
Db 511 ACAAAACGGTAAAAAA 526

Query Match   24.2%; Score 16; DB 12; Length 561;
Best Local Similarity   100.0%; Pred. No. 1.8e+02;
Matches 16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

LOCUS   H1612678   571 bp   mRNA   linear   EST 07-SEP-2001
DEFINITION RH41626.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH41626 5 similar to yin:
            FBan0002913 GO:[integral plasma membrane protein (GO:0005887);
            hydrogen/oligopeptide symporter (GO:0005427); hydrogen/oligopeptide
            symporter (GO:0005427); transporter (GO:0005215); integral plasma
            membrane protein (GO:0005887); hy, mRNA sequence.

```
ACCESSION   BI612678
VERSION     BI612678.1  GI:15508210
KEYWORDS    EST.
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 571)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RH Drosophila EST Project.
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton,M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: RH 416 row: C column: 2
            High quality sequence stop: 505.
FEATURES             Location/Qualifiers
     source          1..571
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH41626"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      152 a    148 c    122 g    149 t
ORIGIN

RESULT 26

Qy      19 GACAAAACGGTAAAAA 34
           ||||||||||||||||
Db     330 GACAAAACGGTAAAAA 345

LOCUS       TA55H11Q      574 bp    DNA     linear    GSS 13-DEC-2000
DEFINITION  T. brucei sheared genomic DNA clone 55h11, reverse sequence,
            genomic survey sequence.
ACCESSION   AL456183
VERSION     AL456183.1  GI:11857054
KEYWORDS    GSS.
SOURCE      Trypanosoma brucei
ORGANISM    Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1  (bases 1 to 574)
  AUTHORS   Hall,N., Bowman,S., Lennard,N.J., Doggett,J., Atkin,R.,
            Chillingworth,C., Ormond,D., Harris,B., El-Sayed,N., Hou,L.,
            Melville,S.E., Rajandream,M.A. and Barrell,B.G.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-2000) Trypanosoma brucei genome sequencing
            project, Sanger Centre, The Wellcome Trust Genome Campus, Hinxton,
            Cambridge CB10 1SA, E-mail: barrell@sanger.ac.uk and
            nh1@sanger.ac.uk
COMMENT     Constructed at the Institute for Genomic Research (TIGR),
```

TA55H11Q/c
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy   Query Match 24.2%; Score 16; DB 10; Length 571;

```
COMMENT     Rockville, MD. Genomic DNA isolated from a cloned population of
            Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically sheared
            to give a tight size distribution (
            4 kb). The v + i method used for the library construction is
            described in detail in Smith, H. and Venter, J.C. (Making small
            insert libraries for whole genome shotgun sequencing projects. In
            Genome Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Barrell, Oxford University Press, 1999).
            Email: nelsayed@tigr.org
            Details of T. brucei sequencing at the Sanger Centre are available
            at http://www.sanger.ac.uk/Projects/T_brucei/.
FEATURES             Location/Qualifiers
     source          1..574
                     /organism="Trypanosoma brucei"
                     /strain="TREU927"
                     /db_xref="taxon:5691"
                     /clone="55h11"
BASE COUNT      184 a    113 c    116 g    161 t
ORIGIN

RESULT 27

Qy      20 ACAAAACGGTAAAAA 35
           |||||||||||||||
Db     193 ACAAAACGGTAAAAA 178

LOCUS       BI565318      580 bp    mRNA    linear    EST 06-SEP-2001
DEFINITION  RH63067.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH63067 5 similar to yin:
            FBan00002913 CG:[integral plasma membrane protein (GO:0005887);
            hydrogen/oligopeptide symporter (GO:0005427); transporter
            symporter (GO:0005427); transporter (GO:0005215); hydrogen/oligopeptide
            membrane protein (GO:0005887), mRNA sequence.
ACCESSION   BI565318
VERSION     BI565318.1  GI:15454510
KEYWORDS    EST.
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RH Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton,M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: RH 630 row: F column: 7
            High quality sequence stop: 515.
FEATURES             Location/Qualifiers
     source          1..580
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH63067"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5 alpha TonA"
```

BI565318/c
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy   Query Match 24.2%; Score 16; DB 12; Length 574;

```
            /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
            BamHI; Library was kindly generated by Piero Carninci at
            the RIKEN. The library was normalized and excised using
            cre recombinase. Plasmid cDNA library."
BASE COUNT     155 a    151 c    124 g    150 t
ORIGIN
Qy      19 GACAAAACGGTAAAAA 34
           |||||||||||||||
Db     340 GACAAAACGGTAAAAA 355

Query Match      24.2%;  Score 16;  DB 10;  Length 580;
Best Local Similarity  100.0%;  Pred. No. 1.3e+02;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

RESULT 28
LOCUS       A14B5800           590 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST244121 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
            cLED4117, mRNA sequence.
ACCESSION   A14B5800
VERSION     A14B5800.1  GI:4481171
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 590)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Ronning,C.M., Ronning,C.M., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
            ,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato carpel tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES             Location/Qualifiers
     source          1..590
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone_lib="tomato ovary, TAMU"
                     /clone="cLED4117"
                     /tissue_type="carpel"
                     /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; cLED - Tomato Carpel EST library. Oligod1-primed and
                     directionally cloned cDNA in vector lamda ZAP II with 5'
                     and 3' ends located at the EcoRI and XhoI sites,
                     respectively."
BASE COUNT     175 a    115 c     92 g    208 t
ORIGIN
Qy      47 CCGTCCTACGAATTCA 62
           ||||||||||||||||
Db     552 CCGTCCTACGAATTCA 567

Query Match      24.2%;  Score 16;  DB 9;  Length 590;
Best Local Similarity  100.0%;  Pred. No. 1.3e+02;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

RESULT 29
LOCUS       B16B0B19           600 bp    mRNA    linear   EST 10-SEP-2001
DEFINITION  RH60131.5prime RH Drosophila melanogaster normalized Head pFlc-1
```

```
            Drosophila melanogaster cDNA clone RH60131 5 similar to yin:
            FBan0002913 GO:[integral plasma membrane protein (GO:0005887);
            hydrogen/oligopeptide symporter (GO:0005427); hydrogen/oligopeptide
            symporter (GO:0005427); transporter (GO:0005215); integral plasma
            membrane protein (GO:0005887), mRNA sequence.
ACCESSION   B16B0B19
VERSION     B16B0B19.1  GI:15533029
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 600)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Bernan,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Paclcb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RH Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: RH.601 row: C  column: 7
            High quality sequence stop: 532.
FEATURES             Location/Qualifiers
     source          1..600
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH60131"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT     157 a    156 c    130 g    155 t    2 others
ORIGIN
Qy      19 GACAAAACGGTAAAAA 34
           |||||||||||||||
Db     340 GACAAAACGGTAAAAA 355

Query Match      24.2%;  Score 16;  DB 10;  Length 600;
Best Local Similarity  100.0%;  Pred. No. 1.3e+02;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

RESULT 30
LOCUS       B156B118           611 bp    mRNA    linear   EST 06-SEP-2001
DEFINITION  RH39495.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH39495 5 similar to yin:
            FBan0002913 GO:[integral plasma membrane protein (GO:0005887);
            hydrogen/oligopeptide symporter (GO:0005427); transport
            symporter (GO:0005427); transporter (GO:0005215); integral plasma
            membrane protein (GO:000588, mRNA sequence.
ACCESSION   B156B118
VERSION     B156B118.1  GI:15457310
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
```

RESULT 31
```
LOCUS       BI237091                617 bp    mRNA    linear   EST 12-JUL-2001
DEFINITION  RE33250 5prime RE Drosophila melanogaster cDNA clone RE33250 5 similar to yin:
            Drosophila melanogaster normalized Embryo pFlc-1
            FBan0002913 'transporter' located on: X 4A1-4A1:; 05/12/2001, mRNA
            sequence.
ACCESSION   BI237091
VERSION     BI237091.1  GI:14705552
KEYWORDS    EST.
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 617)
AUTHORS     Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
TITLE       BDGP/HHMI RE Drosophila EST Project
JOURNAL     Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Fax: 510 486 6798
FEATURES             Location/Qualifiers
     source          1..617
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE33250"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo: Vector: pFlc1: Site_1: XhoI: Site_2:
                     BamHI: Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
                     hit genomic AE003429: arm:X [3604901,3904520]
                     estimated-cyto:3F6-4B1: 05/12/2001
                     Plate: RE.332 row: E column: 2
                     High quality sequence stop: 538.
BASE COUNT      175 a    137 c    134 g    171 t
ORIGIN
```

```
Qy          19 GACAAAACGTAAAAA 34
               ||||||||||||||
Db         340 GACAAAACGGTAAAAA 355
```

```
Query Match            24.2%;  Score 16;  DB 10;  Length 611;
Best Local Similarity  100.0%;                         Pred. No. 1.3e-02;
Matches  16;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

RESULT 32
```
LOCUS       BE442560                629 bp    mRNA    linear   EST 25-JUL-2000
DEFINITION  WHE1103_D05_H09ZS Wheat etiolated seedling root normalized cDNA
            library Triticum aestivum cDNA clone WHE1103_D05_H09, mRNA
            sequence.
ACCESSION   BE442560
VERSION     BE442560.1  GI:9442174
KEYWORDS    EST.
SOURCE      bread wheat.
ORGANISM    Triticum aestivum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Triticum.
REFERENCE   1  (bases 1 to 629)
AUTHORS     Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Hah
            n,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Nguyen,H.T.,
            Rausch,C.J., Seaton,C.L., Tong,J.C. and Zhang,D.
TITLE       The structure and function of the expressed portion of the wheat
            genomes - Normalized root cDNA library
JOURNAL     Unpublished (2000)
COMMENT     Contact: Olin Anderson
            US Department of Agriculture, Agriculture Research Service, Pacific
            West Area, Western Regional Research Center
            800 Buchanan Street, Albany, CA 94710, USA
            Tel: 5105595773
            Fax: 5105595818
            Email: oandersn@pw.usda.gov
            Seq primer: Stratagene SK primer.
            The structure and function of the expressed portion of the wheat
            genomes - Normalized root cDNA library
            Sequence have been trimmed to remove vector sequence and low
            quality sequence with phred score less than 20
FEATURES             Location/Qualifiers
     source          1..629
                     /organism="Triticum aestivum"
                     /cultivar="Chinese Spring"
                     /db_xref="taxon:4565"
                     /clone="WHE1103_D05_H09"
                     /clone_lib="Wheat etiolated seedling root normalized cDNA
                     library"
                     /tissue_type="Root"
                     /dev_stage="Five day old etiolated seedling"
                     /lab_host="E. coli DH10B"
                     /note="Vector: Lambda Uni-ZAP XR, excised phagemid
```

```
Qy          19 GACAAAACGGTAAAAA 34
               ||||||||||||||||
Db         339 GACAAAACGGTAAAAA 354
```

```
Query Match            24.2%;  Score 16;  DB 10;  Length 617;
Best Local Similarity  100.0%;                         Pred. No. 1.3e-02;
Matches  16;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;
```

```
REFERENCE   1  (bases 1 to 611)
AUTHORS     Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
TITLE       BDGP/HHMI RE Drosophila EST Project
JOURNAL     Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Fax: 510 486 6798
            Plate: RH.394 row: H column: 11
            High quality sequence stop: 516.
FEATURES             Location/Qualifiers
     source          1..611
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH39495"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head: Vector: pFlc1: Site_1: XhoI: Site_2:
                     BamHI: Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      161 a    159 c    133 g    158 t
ORIGIN
```

RESULT 44

Qy   7 CTGGATCCGTTCGACA 22
              ||||||||||||||||
Db   245 CTGGATCCGTTCGACA 250

AQ950782/C

Query Match   24.2%; Score 16; DB 10; Length 629;
Best Local Similarity   100.0%; Pred. No. 1.3e+02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

LOCUS   AQ950782   615 bp   DNA   linear   GSS 27-JAN-2000
DEFINITION Sheared DNA-51K2.TF sheared DNA Trypanosoma brucei genomic clone
            Sheared DNA-51K2, DNA sequence.
ACCESSION   AQ950782
VERSION   AQ950782.1 GI:6774047
KEYWORDS   GSS.
SOURCE   Trypanosoma brucei.
ORGANISM   Trypanosoma brucei
            Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae;
            Trypanosoma.
REFERENCE   1 (bases 1 to 615)
AUTHORS   El-Sayed,N., Zhao,S., Zhao,H., Gill,S., Suh,E., Malek,J., Fujii,C.,
            Gerrard,C., Leech,V., de Jong,P., Ulla,E., Melville,S., Donelson,J.,
            Fraser,C., and Adams,M.
TITLE   Determination of clone end sequences from Trypanosoma brucei GUTat
            10.1 sheared DNA library
JOURNAL   Unpublished (1999)
COMMENT   other_GSS: Sheared DNA-51K2.TR
            Contact: Najib M. El-Sayed
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: nelsayed@tigr.org
            Clones are derived from the Trypanosoma brucei GUTat 10.1 sheared
            DNA library constructed at TIGR. Clones will be available for
            distribution through ATCC. Sheared DNA end sequences search page:
            http://www.tigr.org/tdb/mdb/tbdb/.
            Seq primer: M14-Forward
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..615
                     /organism="Trypanosoma brucei"
                     /strain="TREU927/4 GUTat 10.1"
                     /db_xref="taxon:5691"
                     /clone="Sheared DNA-51K2"
                     /clone_lib="Sheared DNA"
                     /note="Vector: pUC18; Site_1: Small; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from a cloned population of
                     Trypanosoma brucei (TREU927/4 GUTat 10.1) was mechanically
                     sheared to give a tight size distribution (approx 2 kb).
                     The v + i method used for the library construction is
                     described in detail in Smith, H.O., and Venter, J.C.
                     (Making small insert libraries for whole genome shotgun
                     sequencing projects. In Genome Sequencing: A Practical
                     Approach, eds. M. Vaudin and B. Harcl), Oxford University
                     Press, 1999)."

Qy   20 ACAAAACGGTAAAAAA 35
              ||||||||||||||||
Db   457 ACAAAACGGTAAAAAA 442

RESULT 34

Qy   20 ACAAAACGGTAAAAAA 35
              ||||||||||||||||
Db   457 ACAAAACGGTAAAAAA 442

BI624045

Query Match   24.2%; Score 16; DB 12; Length 647;
Best Local Similarity   100.0%; Pred. No. 1.3e+02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

LOCUS   BI624045   647 bp   mRNA   linear   EST 07-SEP-2001
DEFINITION RH56430.5prime RH Drosophila melanogaster cDNA clone RH56430 5 similar to yin;
            Drosophila melanogaster cDNA clone RH56430 5 similar to yin:
            FBan0002913 GO:[integral plasma membrane protein (GO:0005887);
            hydrogen/oligopeptide symporter (GO:0005427); hydrogen/oligopeptide
            symporter (GO:0005427); transporter (GO:0005215); integral plasma
            membrane protein (GO:0005887), mRNA sequence.
ACCESSION   BI624045
VERSION   BI624045.1 GI:15519570
KEYWORDS   EST.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 647)
AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson,
            J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S., and Rubin
            ,G.M.
TITLE   BDGP/HHMI RH Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est*fruitfly.berkeley.edu
            Plate: RH.564 row: C column: 6
            High quality sequence stop: 551.
FEATURES   Location/Qualifiers
     source   1..647
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH56430"
                     /clone_lib="RH Drosophila melanogaster normalized head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head; Vector: pFlc); Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     cre recombinase. plasmid cDNA library."

Qy   20 ACAAAACGGTAAAAAA 35
              ||||||||||||||||
Db   457 ACAAAACGGTAAAAAA 442

RESULT 44

AQ950782/C

Query Match   24.2%; Score 16; DB 10; Length 629;
Best Local Similarity   100.0%; Pred. No. 1.3e+02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   7 CTGGATCCGTTCGACA 22
              ||||||||||||||||
Db   245 CTGGATCCGTTCGACA 250

BASE COUNT   161 a   124 c   98 g   252 t
ORIGIN

RESULT 34

BI624045

BASE COUNT   170 a   163 c   142 g   172 t
ORIGIN

Qy   19 GACAAAACGGTAAAAA 34

RESULT 35

```
Db   340 GACAAAACGGTAAAAA 355
         |||||||||||||||||
         AA202447
```

LOCUS   AA202447   649 bp   mRNA   linear   EST 19-APR-2001
DEFINITION LD02644 5prime LD Drosophila melanogaster embryo BlueScript
            Drosophila melanogaster cDNA clone LD02644 5prime, mRNA sequence.
ACCESSION   AA202447
VERSION   AA202447.1 GI:1798197
KEYWORDS   EST.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 649)
AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
TITLE   BDGP/HHMI Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 26 row: D column: 8
            High quality sequence stop: 601.
FEATURES   Location/Qualifiers
     source   1..649
                     /organism="Drosophila melanogaster"
                     /db_xref="BDGP_EST:BDcln002311"
                     /db_xref="taxon:7227"
                     /clone="LD02644"
                     /clone_lib="LD Drosophila melanogaster embryo BlueScript"
                     /sex="male and female"
                     /dev_stage="0 to 24 hours mixed stage embryonic"
                     /lab_host="SOLR"
                     /note="Organ: embryo; Vector: BlueScript SK; Site_1: EcoRI
                     ; Site_2: XhoI; Constructed using Stratagene ZAP-cDNA
                     Synthesis kit. Oligo dT-primed and directionally cloned at
                     EcoRI and XhoI in BlueScript SK(+/-)"
BASE COUNT   169 a   163 c   145 g   172 t
ORIGIN

```
Qy   19 GACAAAACGGTAAAAA 34
        |||||||||||||||||
Db  315 GACAAAACGGTAAAAA 330
```

Query Match   24.2%; Score 16; DB 9; Length 649;
Best Local Similarity   100.0%; Pred. No. 1.1e+02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 36

```
BI229347
|||||||||||||||||
BI229347
```

LOCUS   BI229347   653 bp   mRNA   linear   EST 11-JUL-2001
DEFINITION RE27113.5prime RE Drosophila melanogaster normalized Embryo pflc-1
            Drosophila melanogaster cDNA clone RE27113 5 similar to yin;
            FBan0002913 'transporter' located on: X 4A1-4A1:: 04/12/2001, mRNA
            sequence.
ACCESSION   BI229347
VERSION   BI229347.1 GI:14696611
KEYWORDS   EST.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 653)
AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson
            , J., Champe,M., Chavez,C., Dorsett,V., Farian,D., Frise,E., George
            , R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Colniker,S. and Rubin
            ,G.M.
TITLE   BDGP/HHMI RE Drosophila EST Project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: RE 271 row: B column: 1
            High quality sequence stop: 585.
FEATURES   Location/Qualifiers
     source   1..653
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE27113"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pflc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: embryo; Vector: pflc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT   173 a   165 c   144 g   171 t
ORIGIN

```
Qy   19 GACAAAACGGTAAAAA 34
        |||||||||||||||||
Db  339 GACAAAACGGTAAAAA 354
```

Query Match   24.2%; Score 16; DB 10; Length 653;
Best Local Similarity   100.0%; Pred. No. 1.3e+02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 37

```
AW690071/c
|||||||||||||||||
AW690071
```

LOCUS   NF028A06ST1F1000   660 bp   mRNA   linear   EST 15-JUN-2000
DEFINITION NF028A06ST 5', mRNA sequence.
            Developing stem Medicago truncatula cDNA clone
ACCESSION   AW690071
VERSION   AW690071.1 GI:7564807
KEYWORDS   EST.
SOURCE   barrel medic.
ORGANISM   Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1 (bases 1 to 660)
AUTHORS   He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
            ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
            , R.A.
TITLE   Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula stem library
JOURNAL   Unpublished (2000)
COMMENT   Contact: Dixon RA
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7302
            Fax: 580 221 7380
            Email: radixon@noble.org
            Plate: 028 row: A column: 06
            Insert length: 660 Std Error: 0.00
            Seq primer: TCACACGACAAACAGCTATGAC.
FEATURES   Location/Qualifiers

```
source          1..660
                /organism="Medicago truncatula"
                /db_xref="taxon:3880"
                /clone="NF02BA06ST"
                /clone_lib="developing stem"
                /tissue_type="stem"
                /dev_stage="Pooled developmental"
                /note="Vector: lambda Zap; Contains a mixture of
                internodal stem segments"
BASE COUNT      187 a    160 c    104 g    205 t    4 others
ORIGIN

Qy        2 GACATCTGGATCCCGTT 17
            |||||||||||||||||
Db      497 GACATCTGGATCCCGTT 482


RESULT 48
LOCUS       AZ574452/c
DEFINITION  417PvF10 Pv MBN #10 Plasmodium vivax genomic 3', DNA sequence.
ACCESSION   AZ574452
VERSION     AZ574452.1  GI:11987561
KEYWORDS    GSS.
SOURCE      malaria parasite P. vivax.
  ORGANISM  Plasmodium vivax
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1  (bases 1 to 660)
  AUTHORS   Carlton,J.M.-R. and Dame,J.B.
  TITLE     The Plasmodium vivax and P. berghei gene sequence tag projects
  JOURNAL   Parasitol. Today 16 (10), 409 (2000)
COMMENT     Contact: Dame JB
            Dept. of Pathobiology, College of Veterinary Medicine
            University of Florida
            2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
            Tel: 352 392 4700
            Fax: 352 392 9704
            Email: dame@mail.vetmed.ufl.edu
            Seq primer: M13(-20) forward
            Class: shotgun.
FEATURES             Location/Qualifiers
     source          1..660
                     /organism="Plasmodium vivax"
                     /strain="Salvador I (Collins, W. 1972. J. Parasitol. 69,
                     497-598)"
                     /db_xref="taxon:5855"
                     /clone_lib="Pv MBN #10"
                     /clone="417PvF10"
                     /dev_stage="asexual blood forms"
                     /lab_host="Saimiri boliviensis"
                     /note="Vector: pBluescript SK(+) vector DNA; phagemid
                     excised from lambda ZAP; Site_1: EcoR V; Site_2: EcoR V;
                     Host leukocytes were extracted from P. vivax infected
                     blood using the following methods: first, infected blood
                     was activated by the addition of 0.5 ml of ADP (40mg/ml)
                     per 10 ml blood. Then blood was passed over a column of
                     acid washed 0.1 mm glass beads, then through a Plasmodipur
                     filter, followed by passage through a column of pre-wet
                     Whatman CF11 powder (1:2 ratio volume of blood to CF11),
                     and finally centrifuged through a 50% Percoll density
                     cushion. Purified DNA was digested with mung bean nuclease
                     in the presence of 44% formamide at 50oC as described
                     (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988.
                     Nucleic Acids Research 16:6883-6896). Digested DNA was
                     blunt-ended using T4 DNA Polymerase and size fractionated
                     over a Sepharose CL-2B column. Fractions in the size range
                     500bp-4kb were ligated into the Eco RV site of pBluescript
                     SK(+), and E. coli XL-10 Gold transformed with the
                     ligation mixture."
BASE COUNT      141 a    165 c    90 g    263 t    1 others
```

Best Local Similarity   100.0%;   Pred. No. 1.4e+02;
Query Match   24.2%;   Score 16;   DB 9;   Length 660;
Matches 16;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
ORIGIN

```
ORIGIN

Qy       20 ACAAAAACGGTAAAAAA 45
            ||||||||||||||||
Db      337 ACAAAAACGGTAAAAAA 322
```

Query Match   24.2%;   Score 16;   DB 12;   Length 668;
Best Local Similarity   100.0%;   Pred. No. 1.4e+02;
Matches 16;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
RESULT 39
LOCUS       BI234820
DEFINITION  RE30591.5prime RE Drosophila melanogaster cDNA clone RE30591 5 similar to yid:
            Drosophila melanogaster normalized Embryo pFlc-1
            FBan0002913 'transporter' located on: X 4A1-4A1; 05/12/2001, mRNA
            sequence.
ACCESSION   BI234820
VERSION     BI234820.1  GI:14703272
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 668)
  AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Herman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mundal,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
  TITLE     BDGP/HHMI RE Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: RE5 row: H column: 7
            High quality sequence stop: 641.
FEATURES             Location/Qualifiers
     source          1..668
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RE30591"
                     /clone_lib="RE Drosophila melanogaster normalized Embryo
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="0-24 hours mixed stage embryonic"
                     /lab_host="DH5-alpha TonA"
                     /note="Organism: embryo; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."
BASE COUNT      173 a    166 c    151 g    178 t
ORIGIN
```

Query Match   24.2%;   Score 16;   DB 10;   Length 668;
Best Local Similarity   100.0%;   Pred. No. 1.3e+02;
Matches 16;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

```
Qy       19 GACAAAACGGTAAAAA 34
            ||||||||||||||||
Db      316 GACAAAACGGTAAAAA 351


RESULT 40
LOCUS       AZ572058/c
DEFINITION  302PvH03 Pv MBN #10 Plasmodium vivax genomic 3', DNA sequence.
```

ACCESSION   AZ572058
VERSION   AZ572058.1 GI:13984956
KEYWORDS   GSS.
SOURCE   malaria parasite P. vivax.
ORGANISM   Plasmodium vivax
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1 (bases 1 to 672)
AUTHORS   Carlton,J.M.-R. and Dame,J.B.
TITLE   The Plasmodium vivax and P. berghei gene sequence tag projects
JOURNAL   Parasitol. Today 16 (10), 409 (2000)
COMMENT   Contact: Dame JB
            Dept. of Pathobiology, College of Veterinary Medicine
            University of Florida
            2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
            Tel: 352 392 4700
            Fax: 352 392 9704
            Email: damej@mail.vetmed.ufl.edu
            Seq primer: M13(-20) forward
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..672
                     /organism="Plasmodium vivax"
                     /strain="Salvador I (Collins, W. 1972. J. Parasitol. 69,
                     497-598)."
                     /db_xref="taxon:5855"
                     /clone_lib="Pv MBN #30"
                     /dev_stage="asexual blood forms"
                     /lab_host="Saimiri boliviensis"
                     /note="Vector: pBluescript SK(+) vector DNA, phagemid
                     excised from lambda ZAP; Site_1: EcoR V; Site_2: EcoR V;
                     Host leukocytes were extracted from P. vivax infected
                     blood using the following methods: first, infected blood
                     was activated by the addition of 0.5 ml of ADP (40mg/ml)
                     per 10 ml blood. Then blood was passed over a column of
                     acid washed 0.1 mm glass beads, then through a Plasmodipur
                     filter, followed by passage through a column of pre-wet
                     Whatman CF11 powder (1:2 ratio volume of blood to CF11),
                     and finally centrifuged through a 50% Percoll density
                     cushion. Purified DNA was digested with mung bean nuclease
                     in the presence of 44% formamide at 50oC as described
                     (Vernick, K.D., Imberski, R.B., and McCutchan, T.F. 1988.
                     Nucleic Acids Research 16:6883-6896). Digested DNA was
                     blunt-ended using T4 DNA Polymerase and size fractionated
                     over a Sepharose CL-2B column. Fractions in the size range
                     500bp-4kb were ligated into the Eco RV site of pBluescript
                     SK(+), and E. coli XL-10 Gold transformed with the
                     ligation mixture."

BASE COUNT   140 a   171 c   101 g   259 t   1 others
ORIGIN

RESULT 41
Query Match   24.2%; Score 16; DB 12; Length 672;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   20 ACAAAACGGTAAAAAA 35
              ||||||||||||||||
Db   324 ACAAAACGGTAAAAAA 309

ACCESSION   BI621542
VERSION   BI621542.1 GI:15517067
KEYWORDS   EST.
SOURCE   fruit fly.
ORGANISM   Drosophila melanogaster
LOCUS   BI621542   673 bp   mRNA   linear   EST 07-SEP-2001
DEFINITION RH52894.5prime RH Drosophila melanogaster normalized Head pFlc-1
            Drosophila melanogaster cDNA clone RH52894 5 similar to yin:
            FHan00Q913 GO:[integral plasma membrane protein (GO:0005887);
            hydrogen/oligopeptide symporter (GO:0005427);
            symporter (GO:0005427); transporter (GO:0005215); hydrogen/oligopeptide
            membrane protein (GO:0005887), mRNA sequence.

TITLE   The Plasmodium vivax and P. berghei gene sequence tag projects
JOURNAL   Parasitol. Today 16 (10), 409 (2000)
COMMENT   Contact: Dame JB
            Dept. of Pathobiology, College of Veterinary Medicine
            University of Florida
            2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
            Tel: 352 392 4700
            Fax: 352 392 9704
            Email: damej@mail.vetmed.ufl.edu
            Seq primer: M13(-20) forward
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..673
                     /organism="Plasmodium vivax"
                     /strain="Salvador I (Collins, W. 1972. J. Parasitol. 69,

---

REFERENCE
AUTHORS   Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Herman,B., Carlson
            ,J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Farlan,D., Frise,E., George
            ,R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin
            ,G.M.
TITLE   BDGP/HHMI RH Drosophila EST project
JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: RH_528 row: H column:10
            High quality sequence stop: 565.
FEATURES   Location/Qualifiers
     source   1..673
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="RH52894"
                     /clone_lib="RH Drosophila melanogaster normalized Head
                     pFlc-1"
                     /sex="male and female"
                     /dev_stage="Adult"
                     /lab_host="DH5-alpha TonA"
                     /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
                     BamHI; Library was kindly generated by Piero Carninci at
                     the RIKEN. The library was normalized and excised using
                     Cre recombinase. Plasmid cDNA library."

BASE COUNT   174 a   168 c   152 g   179 t
ORIGIN

RESULT 42
LOCUS   AZ569442
ACCESSION   AZ569442
VERSION   AZ569442.1 GI:13979541
KEYWORDS   GSS.
SOURCE   malaria parasite P. vivax.
ORGANISM   Plasmodium vivax
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE   1 (bases 1 to 679)
AUTHORS   Carlton,J.M.-R. and Dame,J.B.
TITLE   The Plasmodium vivax and P. berghei gene sequence tag projects
JOURNAL   Parasitol. Today 16 (10), 409 (2000)
COMMENT   Contact: Dame JB
            Dept. of Pathobiology, College of Veterinary Medicine
            University of Florida
            2015 SW 23rd Avenue, Bldg 1017, Gainesville, FL 32611, USA
            Tel: 352 392 4700
            Fax: 352 392 9704
            Email: damej@mail.vetmed.ufl.edu
            Seq primer: M13(-20) forward
            Class: shotgun.
FEATURES   Location/Qualifiers
     source   1..679
                     /organism="Plasmodium vivax"
                     /strain="Salvador I (Collins, W. 1972. J. Parasitol. 69,

LOCUS   AZ569442   679 bp   DNA   linear   GSS 15-MAY-2001
DEFINITION 260PvH02 Pv MBN #30 Plasmodium vivax genomic 3'', DNA sequence.

Query Match   24.2%; Score 16; DB 10; Length 673;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches   16; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   19 GACAAAACGGTAAAAA 34
              |||||||||||||||
Db   340 GACAAAACGGTAAAAA 355

RESULT 43

```
BASE COUNT    247 a    116 c    161 g    154 t    1 others
ORIGIN
        497-598)"
        /db_xref "taxon:5855"
        /clone_lib "pv MBN #30"
        /dev_stage "asexual blood forms"
        /lab_host "Saimiri boliviensis"
        /note="Vector: pHluescript SK(+) vector DNA, phagemid
        excised from lambda ZAP; Site_1: EcoR V; Site_2: EcoR V;
        Host leukocytes were extracted from P. vivax infected
        blood using the following methods: first, infected blood
        was activated by the addition of 0.5 ml of ADP (40mg/ml)
        per 10 ml blood. Then blood was passed over a column of
        acid washed 0.1 mm glass beads, then through a plasmodipur
        filter, followed by passage through a column of pre-wet
        Whatman CF11 powder (1:2 ratio volume of blood to CF11),
        and finally centrifuged through a 50% Percoll density
        cushion. Purified DNA was digested with mung bean nuclease
        in the presence of 44% formamide at 50oc as described
        (Verlick, K.D., Imberski, R.B., and McCutchan, T.F. 1988.
        Nucleic Acids Research 16:6883-6896). Digested DNA was
        blunt-ended using T4 DNA Polymerase and size fractionated
        over a Sepharose CL-2B column. Fractions in the size range
        500bp-4kb were ligated into the Eco RV site of pHluescript
        SK(+), and E. coli XL-10 Gold transformed with the
        ligation mixture."
```

```
Qy     20 ACAAAAAGGTAAAAA 45
          |||||||||||||||
Db    217 ACAAAAACGGTAAAAA 242
```

```
RESULT 44
BI582214
BI582214.1  GI:15473636
BI582214

LOCUS       BI582214    681 bp    mRNA    linear    EST 06-SEP-2001
DEFINITION  RH20171.5prime RH Drosophila cDNA clone RH20171 5 similar to yln:
            FBan0002913 GO:[integral plasma membrane protein (GO:0005887);
            hydrogen/oligopeptide symporter (GO:0005427); transporter
            (GO:0005215); integral plasma membrane protein (GO:0005427);
            hydrogen/oligopeptide symporter (GO:0005427); hydrogen/oligopeptide
            symporter (GO:0005215); ty, mRNA sequence.
VERSION
ACCESSION
SOURCE      fruit fly.
            Drosophila melanogaster
KEYWORDS    EST.
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 681)
AUTHORS     Stapleton,M., Brokstein,P., Hong,L., Tyler,D., Berman,B., Carlson,
            J., Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,
            R., Gonzalez,M., Guarin,H., Harris,N., Li,P., Liao,G., Misra,S.,
            Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,
            Phouanenavong,S., Wan,K., Yu,C., Lewis,S.E., Celniker,S. and Rubin,
            G.M.
TITLE       BDGP/HHMI RH Drosophila EST Project
JOURNAL     Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            HHGP
            Lawrence Berkeley National lab
            one Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: RH.201 row: F column:11
            High quality sequence stop: 510.
FEATURES             Location/Qualifiers
     source          1..681
                     /organism="Drosophila melanogaster"
```

```
Query Match    24.2%;    Score 16;    DB 12;    Length 679;
Best Local Similarity   100.0%;    Pred. No. 1.4e+02;
Matches 16;    Conservative 0;    Mismatches 0;    Indels 0;    Gaps 0;
```

RESULT 44

```
BASE COUNT    175 a    171 c    152 g    183 t
ORIGIN
        /db_xref "taxon:7227"
        /clone "RH20171"
        /clone_lib="RH Drosophila melanogaster normalized Head
        pFlc-1"
        /sex "male and female"
        /dev_stage "Adult"
        /note="Organ: head; Vector: pFlc1; Site_1: XhoI; Site_2:
        BamHI; library was kindly generated by Pietro Carninci at
        the RIKEN. The library was normalized and excised using
        Cre recombinase. Plasmid cDNA library."
```

```
Qy     19 GACAAAACGGTAAAAA 34
          ||||||||||||||||
Db    340 GACAAAACGGTAAAAA 355
```

```
RESULT 44
BB432863
BB432863
BB432863.2  GI:16424253

LOCUS       BB432863    690 bp    mRNA    linear    EST 25-OCT-2001
DEFINITION  BB432863 RIKEN full-length enriched, adult male hippocampus Mus
            musculus cDNA clone C630048D20 3', mRNA sequence.
ACCESSION   BB432863
VERSION     BB432863.2  GI:16424253
KEYWORDS    EST.
SOURCE      house mouse.
ORGANISM    Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 690)
AUTHORS     Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Takeda,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
TITLE       RIKEN Mouse ESTs (Arakawa,T. et al. 2001)
JOURNAL     Unpublished (2001)
COMMENT     On Jul 18, 2000 this sequence version replaced gi:9272590.
            Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
            Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. 10 (10), 1617-1630 (2000)
            wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,M., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
            RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res.
            10 (11), 1757-1771 (2000)
            Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
            Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
            Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
```

```
Query Match    24.2%;    Score 16;    DB 10;    Length 681;
Best Local Similarity   100.0%;    Pred. No. 1.4e+02;
Matches 16;    Conservative 0;    Mismatches 0;    Indels 0;    Gaps 0;
```

Left record

```
REFERENCE   K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
  TITLE     Computational Analysis of Full-length Mouse cDNAs Compared with
            Human Genome Sequences
  JOURNAL   Mamm. Genome. 12, 673-677 (2001)
            Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
  COMMENT   cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in Riken
            Division of Experimental Animal Research in RIKEN contributed to
            prepare mouse tissues;

FEATURES          Location/Qualifiers
  source          1..690
                  /organism="Mus musculus"
                  /strain="C57BL/6J"
                  /db_xref="taxon:10090"
                  /clone="C630048D20"
                  /clone_lib="RIKEN full-length enriched, adult male
                  hippocampus"
                  /sex="male"
                  /tissue_type="hippocampus"
                  /dev_stage="adult"
                  /lab_host="SOLR"
                  /note="Site_1: XhoI; Site_2: BamHI; cDNA library was
                  prepared and sequenced in Mouse Genome Encyclopedia
                  Project of Genome Exploration Research Group in RIKEN,
                  Genomic Sciences Center and Genome Science Laboratory in
                  RIKEN, Division of Experimental Animal Research in Riken
                  contributed to prepare mouse tissues. 1st strand cDNA was
                  primed with a primer [5'
                  GAGAGAGAGAGGATCCAAGAGCTCTTTTTTTTTTTTTTVN 3'], cDNA was
                  prepared by using trehalose thermo-activated reverse
                  transcriptase and subsequently enriched for full-length by
                  cap-trapper. Second strand cDNA was prepared with the
                  primer adapter of sequence [5'
                  GAGAGAGAGATTCTCGAGTTTAATTAAATTAATCCCCCCCCCC 3']"

BASE COUNT      267 a    123 c    104 g    196 t
ORIGIN

Query Match             24.28%;  Score 16;  DB 9;  Length 690;
Best Local Similarity  100.0%;   Pred. No. 1.3e+02;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     19 GACAAAACGGTAAAAA 34
          ||||||||||||||||
Db    182 GACAAAACGGTAAAAA 197
```

```
RESULT 45
HI954699/c

LOCUS       HVSMEm0019F17f     694 bp    mRNA    linear   EST 19-OCT-2001
DEFINITION  HVcDNA0014 (Blumeria infected) Hordeum vulgare green seedling EST
            library  Hordeum vulgare cDNA clone HVSMEm0019F17t, mRNA sequence.
ACCESSION   HI954699
VERSION     HI954699.1  GI:16300423
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1 (bases 1 to 694)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Chin,A., Begum,D.,
            Frisch,D., Atkins,M., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons
            ,J., Oates,R. and Main,D.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics: Blumeria infected Morex (compatible) seedling
            cDNA library
  JOURNAL   Unpublished (2001)
  COMMENT   Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
```

Right record

```
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hq bases = 171
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence start: 20
            High quality sequence stop: 242.

FEATURES          Location/Qualifiers
  source          1..694
                  /organism="Hordeum vulgare"
                  /cultivar="Morex"
                  /db_xref="taxon:4513"
                  /clone_lib="HVSMEm0019F17t"
                  /clone="HVcDNA0014 (Blumeria infected)"
                  /tissue_type="green seedling leaf"
                  /lab_host="TJC121"
                  /note="Vector: pBluescript SK(); Site_1: EcoRI; Site_2:
                  XhoI; Morex (mla) plants were greenhouse grown in the R
                  Wise lab at Iowa State University, Ames, IA; 7 day old
                  green seedlings were infected with isolate 5874 of
                  Blumeria graminis f. sp. hordei, and leaves were harvested
                  24, 48 and 72 hr post-inoculation and snap frozen (Wise).
                  In the TJ Close lab at the University of California,
                  Riverside, total RNA was prepared from each sample pool,
                  equal quantities of all three RNA pools were combined,
                  poly(A) RNA was purified from the mixture, one primary
                  unamplified cDNA library was made, and 1 million pfu were
                  in vivo excised to give pBluescript SK(-) cDNA phagemids
                  (chin). Phagemids were plated and picked at the Clemson
                  University Genomics Institute (CUGI) (Begum, Palmer,
                  Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
                  sequencing and sequence analysis were performed at CUGI
                  (Wing, Yu, Frisch, Henry, Simmons, Oates, Rambo, Main).
                  The sequence has been trimmed to remove vector sequence
                  and contains a minimum of 100 bases of phred value 20 or
                  above. For more details on library preparation and
                  sequence analysis see
                  http://www.genome.clemson.edu/projects/barley. To order
                  this clone see http://www.genome.clemson.edu/orders Also
                  see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                  Genetically and physically anchored EST resources for
                  barley genomics. Barley Genetics Newsletter 31:29-30.
                  (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

BASE COUNT      216 a    171 c    128 g    179 t
ORIGIN

Query Match             24.28%;  Score 16;  DB 10;  Length 694;
Best Local Similarity  100.0%;   Pred. No. 1.3e+02;
Matches  16;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy     45 TACCGTCCTACGAATT 60
          ||||||||||||||||
Db     17 TACCGTCCTACGAATT 2

Search completed: November 5, 2002, 10:51:56
Job time : 798.981 secs
```
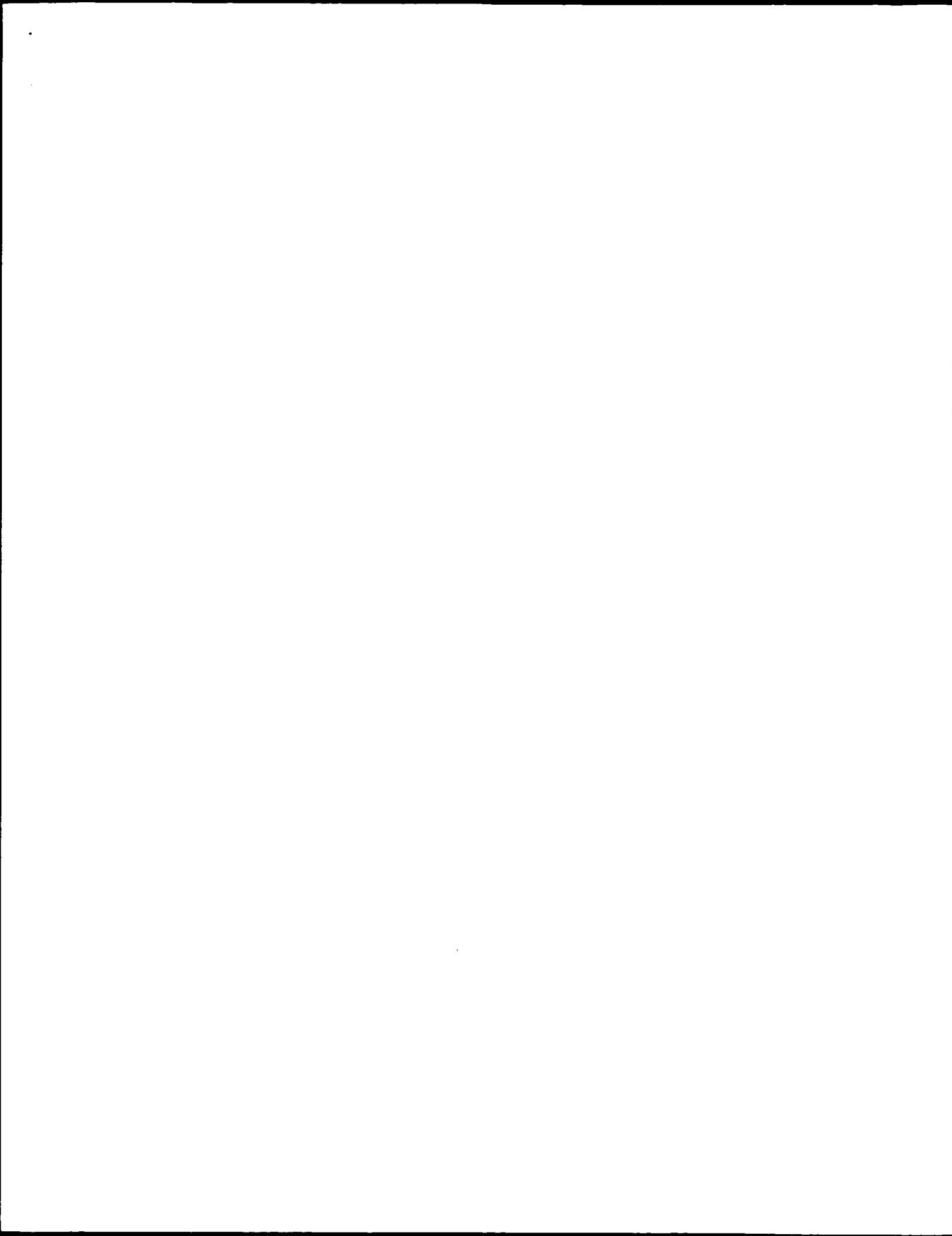
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

GM nucleic - nucleic search, using sw model

Run On: November 5, 2002, 05:28:09 ; Search time 800.027 seconds
(without alignments)
1113.461 Million cell updates/sec

Title: US-09-766-113-2
Perfect score: 66
Sequence: 1 tgaaatctgaatccgttca..........ccgtcctacgaattcagctg 66

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 13746207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Post-processing: Minimum Match 0%
Minimum DB seq length: 0   Maximum Match 100%
Maximum DB seq length: 200000000
Listing first 45 summaries

Database :   EST:*
    1: em_estba:*
    2: em_esthum:*
    3: em_estln:*
    4: em_estmu:*
    5: em_estov:*
    6: em_estpl:*
    7: em_estro:*
    8: em_htc:*
    9: gb_est1:*
   10: gb_est2:*
   11: gb_htc:*
   12: gb_gss:*
   13: em_gss_hum:*
   14: em_gss_iuv:*
   15: em_gss_plu:*
   16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	27.6	41.8	600	12	FR0003332	Z87115 F.rubripes
2	25.4	38.5	382	10	BE318413	BE318413 NF038B09L
3	25.4	38.5	447	9	AW690926	AW690926 NF034F11S
4	25.4	38.5	474	9	AL366452	AL366452 MtBA08A07
5	25.4	38.5	510	9	AA660506	AA660506 00392 MLR
6	25.4	38.5	587	9	AW693673	AW693673 NF067A03S
7	25.4	38.5	587	10	BI308971	BI308971 EST530381
8	25.4	38.5	645	10	BF648610	BF648610 NF049E12E
9	25.4	38.5	651	9	AW696944	AW696944 NF112E09S
10	25.4	38.5	661	9	AW685420	AW685420 NF029B07N
11	25.4	38.5	680	10	BF642214	BF642214 NF067C04I
12	25.4	38.5	688	9	AW689706	AW689706 NF023D11S
13	25.4	38.5	777	10	BG645754	BG645754 EST507373
14	25.4	38.2	459	9	AI260450	AI260450 LP043321.5
15	25.2	38.2	521	9	AA698485	AA698485 HL04775.5
16	25.2	38.2	763	12	BH274021	BH274021 CH230-108
17	25	37.9	877	12	CNS01ZSO	AL174681 Tetraodon

	Score	Query Match	length	DB	ID	Description
c 18	25	37.9	973	12	CNS02SOM	AL212359 Tetraodon
c 19	24.8	37.6	617	12	BH181573	BH181573 019_G_19-
c 20	24.8	37.6	617	12	CNS07N61	AL618523 T3 end of
c 21	24.8	37.6	672	12	BH335229	BH335229 CH230-5Jl
c 22	24.8	37.6	815	12	BH453604	BH453604 ROGNE77TR
c 23	24.6	37.3	379	9	AV656143	AV656143 AV656143
c 24	24.6	37.3	431	12	AQ685562	AQ685562 HS_552H_B
c 25	24.6	37.3	447	12	AZ149975	AZ149975 SP_0005_H
c 26	24.6	37.3	453	12	AA252561	AA252561 zp87h06.r
c 27	24.6	37.3	624	12	AZ185993	AZ185993 SP_1006_T
c 28	24.4	37.0	636	9	AV398704	AV398704 AV398704
c 29	24.4	37.0	665	10	BI891166	BI891166 ZF647-3-0
c 30	24.4	37.0	731	9	AW355423	AW355423 S47A12 AG
c 31	24.4	37.0	751	9	AV398244	AV398244 AV398244
c 32	24.4	37.0	771	9	AU003361	AU003361 AU003361
c 33	24.4	37.0	834	9	AV400516	AV400516 AV400516
c 34	24.4	37.0	526	12	AZ152723	AZ152723 SP_0028_A
c 35	24.2	36.7	556	10	BM094649	BM094649 Saj18q07.
c 36	24.2	36.7	570	10	BI787066	BI787066 Sai56h10
c 37	24.2	36.7	701	10	HH535766	HH535766 H0O1L41TR
c 38	24.2	36.7	703	12	AG090665	AG090665 Pan trogl
c 39	24.2	36.7	717	12	AG090665	AG090665 Pan trogl
c 40	24	36.4	766	12	AZ193011	AZ193011 SP_1022_B
c 41	23.8	36.1	281	9	AQ578341	AQ578341 nbxb0092N
c 42	23.8	36.1	302	10	AW945199	AW945199 PM4 EN000
c 43	23.8	36.1	340	10	BM357468	BM357468 N161V-B3
c 44	23.8	36.1	357	12	AZ314502	AZ314502 1M003I0I4
c 45	23.8	36.1	643	10	BI887279	BI887279 ZF637-1-0

ALIGNMENTS

RESULT 1
FR0003332

LOCUS   FR0003332   600 bp   DNA   linear   GSS 27-FEB-1997
DEFINITION F.rubripes GSS sequence, clone 022F10bE3, genomic survey sequence.
ACCESSION   Z87115
VERSION   Z87115.1 GI:1184027
KEYWORDS   GSS; genome survey sequence.
SOURCE   Takifugu rubripes
  ORGANISM Takifugu rubripes
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Takifugu.
REFERENCE   1 (bases 1 to 600)
  AUTHORS   Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrania,Y.,
            Williams,G. and Brenner,S.
  TITLE   Direct Submission
  JOURNAL   Submitted (18-FEB-1997) MRC Human Genome Mapping Project Resource
            Centre Hinxton, Cambridge, CB10 1SB, Email: biohelp@hamp.mrc.ac.uk
COMMENT   V_type: phagemid
            PRIMER: KS
            DESCR:
            One pass dye-terminator sequencing of cosmid cloned genomic
            sequence.
FEATURES   Location/Qualifiers
     source   1..600
                     /organism="Takifugu rubripes"
                     /db_xref="taxon:31033"
                     /clone_lib="cosmid 022F10"
                     /clone="022F10bE3"
BASE COUNT   156 a   135 c   141 g   136 t   32 others
ORIGIN

Query Match   41.8%; Score 27.6; DB 12; Length 600;
Best Local Similarity   66.1%;
Matches   39; Conservative   0; Mismatches   20; Indels   0; Gaps   0;

Qy 6 TCTGGATCCGTTCGACAAAACGGTAAAAACGGTAGAATTACCGTCCTACGAATTCAGC 64
        |||||||||||| || |||||| ||| ||| | ||||| ||| | || ||| || | ||

Db   459 TCTGATCCGCTCTCACAACAGCGTGTGATAAAACGGGAAAATTATTNTTCAGTCAAACCCGC 417

RESULT 2

LOCUS   BE318414   882 bp   mRNA   linear   EST 21-DEC-2000
DEFINITION   NF038H09LF1F1074 developing leaf Medicago truncatula cDNA clone
             NF038H09LF 5', mRNA sequence.
ACCESSION   BE318414
VERSION   BE318414.2 GI:11962115
KEYWORDS   EST.
SOURCE   barrel medic.
ORGANISM   Medicago truncatula
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
             Medicago.
REFERENCE   1 (bases 1 to 882)
AUTHORS   Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
             Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLE   Expressed Sequence Tags from the Samuel Roberts Noble Foundation
             Medicago truncatula leaf library
JOURNAL   Unpublished (2000)
COMMENT   On Jul 14, 2000 this sequence version replaced gi:9192190.
             Contact: May GD
             Plant Biology Division
             The Samuel Roberts Noble Foundation
             2510 Sam Noble Parkway, Ardmore, OK 74402, USA
             Tel: 580 221 7491
             Fax: 580 221 7380
             Email: gdmay@noble.org
             Medicago Genome Initiative accession: Mt1.S:24262
             Insert Length: 762   Std Error: 0.00
             Plate: 038   row: B   column: 09
             Seq primer: TCACACACAAACAGCTATGAC.
FEATURES   Location/Qualifiers
     source   1..882
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF038H09LF"
                     /clone_lib="developing leaf"
                     /tissue_type="leaf"
                     /dev_stage="Pooled developmental"
                     /note="Vector: Lambda Zap; Contains a mixture of very
                     young, developing, mature and senescing leaves."
BASE COUNT   110 a   82 c   95 g   95 t
ORIGIN
Qy   4 GATCTCGATCCGCTTCGACAAAAACGGTAAAAAAGCGGTAGATTACCGTCTCA 54
            ||||||||||||| |||| |||| ||||||||| | |||| ||||||||||||
Db   57 GATTCTCGTTCCGTTCGATAAAAACCGATTGCTAAAACCATCTCTTATTCTCCTA 107

RESULT 4

LOCUS   AW690926   447 bp   mRNA   linear   EST 15-JUN-2000
DEFINITION   NF034F11ST1F1000 developing stem Medicago truncatula cDNA clone
             NF034F11ST 5', mRNA sequence.
ACCESSION   AW690926
VERSION   AW690926.1 GI:7565585
KEYWORDS   EST.
SOURCE   barrel medic.
ORGANISM   Medicago truncatula
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
             Medicago.
REFERENCE   1 (bases 1 to 447)
AUTHORS   He,X.Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell

Query Match   38.5%;   Score 25.4;   DB 10;   Length 882;
Best Local Similarity 68.6%;   Pred. No. 99;
Matches   45;   Conservative   0;   Mismatches   16;   Indels   0;   Gaps   0;

TITLE   Expressed Sequence Tags from the Samuel Roberts Noble Foundation
             Medicago truncatula stem library
             ,R.A.
JOURNAL   Unpublished (2000)
COMMENT   Contact: Dixon RA
             Plant Biology Division
             The Samuel Roberts Noble Foundation
             2510 Sam Noble Parkway, Ardmore, OK 74402, USA
             Tel: 580 221 7302
             Fax: 580 221 7380
             Email: radixon@noble.org
             Insert Length: 447   Std Error: 0.00
             Plate: 034   row: F   column: 11
             Seq primer: TCACACACAAACAGCTATGAC.
FEATURES   Location/Qualifiers
     source   1..447
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF034F11ST"
                     /clone_lib="developing stem"
                     /tissue_type="stem"
                     /dev_stage="Pooled developmental"
                     /note="Vector: Lambda Zap; Contains a mixture of
                     internodal stem segments"
BASE COUNT   128 a   93 c   110 g   115 t   1 others
ORIGIN
Qy   4 GATCTCGATCCGCTTCGACAAAAACGGTAAAAAAGCGGTAGATTACCGTCTCA 54
            ||||||||||||| |||| |||| ||||||||| | |||| ||||||||||||
Db   57 GATTCTCGTTCCGTTCGATAAAAACCGATTGCTAAAACCGATCTCTTATTCTCCTA 107

Query Match   38.5%;   Score 25.4;   DB 9;   Length 447;
Best Local Similarity 68.6%;   Pred. No. 1e-02;
Matches   35;   Conservative   0;   Mismatches   16;   Indels   0;   Gaps   0;

RESULT 4

LOCUS   AL366452
DEFINITION   MtBA08A07F1 MtBA Medicago truncatula cDNA clone MtBA08A07 T3, mRNA
             sequence.
ACCESSION   AL366452
VERSION   AL366452.1 GI:9666205
KEYWORDS   EST.
SOURCE   barrel medic.
ORGANISM   Medicago truncatula
             Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
             Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
             Medicago.
REFERENCE   1 (bases 1 to 474)
AUTHORS   Journet,E.P., Crespeau,H., van-Tuinen,D., Gouzy,J., Jalllon,O.,
             Niebel,A., Carreau,V., Chataquier,O., Kahn,D., Gianinazzi-Pearson
             ,V. and Gamas,P.
TITLE   Medicago truncatula ESTs from nitrogen-starved roots
JOURNAL   Unpublished (2000)
COMMENT   Contact: Genoscope
             Genoscope - Centre National de Sequencage
             BP 191 91006 EVRY cedex - France
             Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
             Contact: Pascal Gamas and Etienne-Pascal Journet, Laboratoire de
             Biologie Moleculaire des Relations Plante's-Microorganismes,
             CNRS-INRA, BP 27 31326 Castanet-Tolosan Cedex, France (Email :
             Mt-est@toulouse.inra.fr Website :
             http://sequence.toulouse.inra.fr/Mtruncatula.html).
FEATURES   Location/Qualifiers
     source   1..474
                     /organism="Medicago truncatula"
                     /cultivar="Jemalong"
                     /db_xref="taxon:3880"
                     /clone="MtBA08A07"
                     /clone_lib="MtBA"

LOCUS   MtBA08A07F1 MtBA Medicago truncatula cDNA clone MtBA08A07 T3, mRNA
             474 bp   mRNA   linear   EST 04-AUG-2000

Query Match   38.5%;   Score 25.4;   DB 9;   Length 447;
Best Local Similarity 68.6%;   Pred. No. 1e-02;
Matches   35;   Conservative   0;   Mismatches   16;   Indels   0;   Gaps   0;
Qy   4 GATCTCGATCCGCTTCGACAAAAACGGTAAAAAAGCGGTAGATTACCGTCTCA 54
            ||||||||||||| |||| |||| ||||||||| | |||| ||||||||||||
Db   57 GATTCTCGTTCCGTTCGATAAAAACCGATTGCTAAAACCGATCTCTTATTCTCCTA 107

,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon

/tissue_type="root tips"
/dev_stage="harvested after 3 days of N-starvation"
/note="Vector: pBluescript pSK; Site_1: EcoRI; Site_2:
XhoI. Plants were grown in an aeroponic chamber for 14
days on nitrogen-rich medium followed by 3 days on N-free
medium. RNA was extracted from root tips (1-3 cm). cDNA
was prepared from polyA+ enriched RNA. The cDNA was
directionally ligated into Uni-zapXR vector from
Stratagene and packaged using Gigapack Gold packaging
extracts. Plasmids containing cDNA inserts were
mass-excised from phage stocks using ExAssit helper phage
and propagated in SOLR cells. Clone ordering and
sequencing was performed by the Centre National de
Sequencage (Genoscope, Evry, France)."

BASE COUNT   137 a   92 c   123 g   122 t
ORIGIN

Query Match   38.5%;   Score 25.4;   DB 9;   Length 474;
Best Local Similarity   68.6%;   Pred. No. 1e-02;
Matches 35;   Conservative   0;   Mismatches 16;   Indels 0;   Gaps 0;

Qy   4 GATCTGGATCCGTTCCACAAAAACGGTAAAAAAGCCGGTAGAATTACCGTCCTA 54
     ||||||||||||||||| ||| |||| ||||| || ||||| |||||||||||
Db   31 GATTCTGTTCCGTTCGATAAAACCGATTGTAAAACGATGTCTTATCGTCCTA 81

RESULT 5
LOCUS   AA660506   510 bp   mRNA   linear   EST 08-MAR-2000
DEFINITION   00392 MtRHE Medicago truncatula cDNA 5' similar to importin alpha-1
     ; karyopherin alpha-1; nuclear import protein; mRNA sequence.
ACCESSION   AA660506
VERSION   AA660506.1   GI:2604550
KEYWORDS   EST.
SOURCE   barrel medic.
  ORGANISM   Medicago truncatula
     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
     Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
     Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
     Medicago.
REFERENCE   1 (bases 1 to 510)
AUTHORS   Covitz,P.A., Smith,L.S. and Long,S.R.
TITLE   Expressed sequence tags from a root-hair-enriched medicago
     truncatula cDNA library
JOURNAL   Plant Physiol. 117 (4), 1325-1332 (1998)
COMMENT   Contact: Long SR
     Department of Biological Sciences and Howard Hughes Medical
     Institute
     Stanford University
     Gilbert Biology, Stanford, CA 94305-5020, USA
     Tel: 650 723 3232
     Fax: 650 725 8309
     Email: fa.srl@forsythe.stanford.edu
     Seq primer: T3.
FEATURES   Location/Qualifiers
  source   1..510
     /organism="Medicago truncatula"
     /cultivar="Jemalong"
     /db_xref="taxon:3880"
     /clone_lib="MtRHE"
     /tissue_type="Root hairs & tips"
     /dev_stage="2-3 day old seedlings"
     /note="Organ: Root; Vector: pBK-CMV; Site_1: EcoRI;
     Site_2: XhoI; cDNA was synthesized from a pooled mRNA prep
     from elongating root hairs (30% w/w) and 2-3cm root tips
     (70% w/w). XhoI-oligo-dT linker-primer and EcoRI
     adaptors were used. cDNAs was cloned unidirectionally
     into lambda ZAP Express (Stratagene), amplified, and
     mass-excised into pBK-CMV vector plasmids. More
     information is available at http://bio-5RL8.stanford.edu."

BASE COUNT   139 a   105 c   123 g   135 t   8 others
ORIGIN

Query Match   38.5%;   Score 25.4;   DB 9;   Length 510;
Best Local Similarity   68.6%;   Pred. No. 1e-02;
Matches 35;   Conservative   0;   Mismatches 16;   Indels 0;   Gaps 0;

Qy   4 GATCTGGATCCGTTCGACAAAAACGGTAAAAAAGCGGTAGATTACCGTCCTA 54
     ||| ||||||||||| ||| |||| || ||||| || ||||| |||||||||||
Db   43 GATTCTGTTCCGTTCGATAAAACCGATTGTAAAACGATGTCTTATCGTCCTA 93

RESULT 6
LOCUS   AW693673   587 bp   mRNA   linear   EST 20-DEC-2000
DEFINITION   NF067A03STIFI019 Developing stem Medicago truncatula cDNA clone
     NF067A03ST 5', mRNA sequence.
ACCESSION   AW693673
VERSION   AW693673.2   GI:11933778
KEYWORDS   EST.
SOURCE   barrel medic.
  ORGANISM   Medicago truncatula
     Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
     Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
     Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
     Medicago.
REFERENCE   1 (bases 1 to 587)
AUTHORS   He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
     ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
     ,R.A.
TITLE   Expressed Sequence Tags from the Samuel Roberts Noble Foundation
     Medicago truncatula stem library
JOURNAL   Unpublished (2000)
COMMENT   On Apr 14, 2000 this sequence version replaced gi:7568333.
     Contact: Dixon RA
     Plant Biology Division
     The Samuel Roberts Noble Foundation
     2510 Sam Noble Parkway, Ardmore, OK 73402, USA
     Tel: 580 221 7302
     Fax: 580 221 7380
     Email: radixon@noble.org
     Insert Length: 669   Std Error: 0.00
     Plate: 067   row: A   column: 03
     Seq primer: TCACACAGAAACAGCTATGAC.
FEATURES   Location/Qualifiers
  source   1..587
     /organism="Medicago truncatula"
     /db_xref="taxon:3880"
     /clone="NF067A03ST"
     /clone_lib="Developing stem"
     /tissue_type="stem"
     /dev_stage="Pooled developmental"
     /note="Vector: Lambda Zap; Contains a mixture of
     internodal stem segments"

BASE COUNT   165 a   118 c   146 g   158 t
ORIGIN

Query Match   38.5%;   Score 25.4;   DB 9;   Length 587;
Best Local Similarity   68.6%;   Pred. No. 1e-02;
Matches 35;   Conservative   0;   Mismatches 16;   Indels 0;   Gaps 0;

Qy   4 GATCTGGATCCCGTTCGACAAAAACGGTAAAAAAGCGGTAGATTACCGTCCTA 54
     ||| ||||||||||||| ||| |||| ||| ||||| || ||||| ||||||||||
Db   57 GATTCTGTTCCGTTCGATAAAACCGATTGTAAAACGATGTCTTATCGTCCTA 107

RESULT 7
LOCUS   BI308971   587 bp   mRNA   linear   EST 20-JUL-2001
DEFINITION   EST510381 GPOD Medicago truncatula cDNA clone pGPOD-10H6 5' end,
     mRNA sequence.
ACCESSION   BI308971
VERSION   BI308971.1   GI:14982298
KEYWORDS   EST.
SOURCE   barrel medic.
  ORGANISM   Medicago truncatula

REFERENCE   1 (bases 1 to 587)
AUTHORS   Grusak,M.A., Samac,D.A., Town,C.D., Van Aken,S., Utterback,T., Cho
              ,J. and Fraser,C.M.
TITLE   ESTs from developing reproductive tissues of Medicago truncatula
JOURNAL   Unpublished (2001)
COMMENT   Contact: Michael A. Grusak
              USDA/ARS Children's Nutrition Research Center
              Baylor College of Medicine
              1100 Bates Street, Houston, TX 77030-2600, USA
              Tel: 713-798-7044
              Fax: 713-798-7078
              Email: mgrusak@bcm.tmc.edu
              H49582cc
              TIGR sequence name: MtGAZ27TK
              More information is available at: www.medicago.org
              Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES   Location/Qualifiers
     source   1..587
                     /organism="Medicago truncatula"
                     /cultivar="A17"
                     /db_xref="taxon:3880"
                     /clone="pGPOD-10E6"
                     /clone_lib="GPOD"
                     /tissue_type="immature pod walls"
                     /dev_stage="immature pods, ranging in age from 15 to 30
                     days after pollination"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; Immature pods, ranging in age from 15 to 30 days
                     after pollination, were collected from greenhouse-grown
                     plants. At harvest, seeds were removed from pods and
                     isolated pod walls were collected and immediately frozen
                     in liquid nitrogen. Pod walls were pooled for mRNA
                     extraction. cDNA was prepared from polyA+ enriched RNA.
                     The cDNA was directionally ligated into the Unizap XR
                     vector from Stratagene and packaged using Gigapack III
                     Gold packaging extracts. Plasmids containing cDNA inserts
                     were excised from the recombinant lambda-Zap phage using
                     Ex assist helper phage and propagated in XLOLR cells."
BASE COUNT   165 a   117 c   147 g   158 t
ORIGIN

RESULT 8
BF648610
LOCUS   BF648610   645 bp   mRNA   linear   EST 20-DEC-2000
DEFINITION NF049E12QC1F1098 Elicited cell culture Medicago truncatula cDNA
            clone NF049E12EC 5', mRNA sequence.
ACCESSION   BF648610
VERSION   BF648610.1 GI:11913740
KEYWORDS   EST.
SOURCE   barrel medic.
ORGANISM   Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1 (bases 1 to 645)
AUTHORS   Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
TITLE   Expressed Sequence Tags from the Samuel Roberts Noble Foundation -
            Center for Medicago Genomics Research

Query Match   48.5%; Score 25.4; DB 10; Length 587;
Best Local Similarity   68.6%; Pred. No. 1e-02;
Matches   35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy   4 GATCTCGATCTCCTTCCACAAAAACCGTAAAAAAACGCGTAGATATTACCGTCCTA 54
           ||| |||||||||| | || || ||| |||||||| ||| ||||| ||||||||||
Db   57 GATTCTGTTCCGTTCGCATAAAACCGATTGTAAAACCCATGTCTTATCGTCCTA 107

JOURNAL   Unpublished (2000)
COMMENT   Contact: Dixon RA
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 74402, USA
            Tel: 580 221 7302
            Fax: 580 221 7380
            Email: radixon@noble.org
            Clone="NF049E12EC"
            Insert Length: 645   Std Error: 0.00
            Plate: 049 row: E column: 12
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES   Location/Qualifiers
     source   1..645
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF049E12EC"
                     /clone_lib="Elicited cell culture"
                     /tissue_type="cell cultures derived from root tissues"
                     /dev_stage="cell suspensions were subcultured every 14
                     days. Cells were induced six days after subculture"
                     /note="Vector: lambda Zap; Cells were induced with yeast
                     cell wall extracts equivalent to 50ug/ml glucose in the
                     final concentration. Samples were taken at 0.5, 1, 12 and
                     24 hours after induction. Equal amounts of RNA from each
                     time point were pooled and used for mRNA isolation."
BASE COUNT   174 a   140 c   172 t   5 others
ORIGIN

RESULT 9
AW696944
AW696944
LOCUS   AW696944   651 bp   mRNA   linear   EST 15-JUN-2000
DEFINITION NF112E09ST1F1070 Developing stem Medicago truncatula cDNA clone
            NF112E09ST 5', mRNA sequence.
ACCESSION   AW696944
VERSION   AW696944.1 GI:7571706
KEYWORDS   EST.
SOURCE   barrel medic.
ORGANISM   Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1 (bases 1 to 651)
AUTHORS   He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
            ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
            ,R.A.
TITLE   Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL   Unpublished (2000)
COMMENT   Contact: Dixon RA
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 74402, USA
            Tel: 580 221 7302
            Fax: 580 221 7380
            Email: radixon@noble.org
            Insert Length: 651   Std Error: 0.00
            Plate: 112 row: E column: 09
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES   Location/Qualifiers
     source   1..651
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF112E09ST"

Query Match   48.5%; Score 25.4; DB 10; Length 645;
Best Local Similarity   68.6%; Pred. No. 1e-02;
Matches   35; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

Qy   4 GATCTCGGATCCGTTCGACAAAAACGGTAAAAAAAGCGGTAGATATTACCGTCCTA 54
           ||| ||||||||||| | || || ||| ||||||||| || ||||| |||||||||
Db   57 GATTCTGTTCCGTTCGCATAAAACCGATTCTAAAACCCATGTCTTATCGTCCTA 107

Left section

```
                     /clone_lib="Developing stem"
                     /tissue_type="stem"
                     /dev_stage="Pooled developmental"
                     /note="Vector: Lambda Zap; Contains a mixture of
                     internodal stem segments"
BASE COUNT    178 a    130 c    167 g    176 t
ORIGIN
Query Match    38.5%;  Score 25.4;  DB 9;  Length 651;
Best Local Similarity   68.6%;  Pred. No. 1e-02;
Matches   35;  Conservative   0;  Mismatches   16;  Indels   0;  Gaps   0;

Qy    4 GATCTCGGATCCGTTCCACAAAACCCGGTAAAAAAGCCGTAGAITACCGTCCTA 54
         |||||||||||| ||||||| ||||||||||||||||| |||| |||||||||
Db   58 GATTCTGTTCCGTTCGATAAAACCGATTGTAAAAACGATGTCTTATCGTCCTA 108

RESULT 10
LOCUS       AW685420                661 bp    mRNA        linear    EST 15-JUN-2000
DEFINITION  NF029B07NR1F1000 Nodulated root Medicago truncatula cDNA clone
            NF029B07NR 5', mRNA sequence.
ACCESSION   AW685420
VERSION     AW685420.1  GI:7560156
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 661)
  AUTHORS   Watson,B.S., Shin,H.-S., Lopez-Meyer,M., Scott,A.D., Harris,A.R.,
            Gonzales,R.A., Bell,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May
            ,G.D. and Paiva,N.L.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula nodulated root library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Paiva NL
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7317
            Fax: 580 221 7380
            Email: nlpaiva@noble.org
            Insert Length: 661  Std Error: 0.00
            Plate: 029  row: B  column: 07
            Seq primer: TCACACAGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1..661
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF029B07NR"
                     /clone_lib="Nodulated root"
                     /tissue_type="root"
                     /dev_stage="Four-week old Rhizobium
                     meliloti-inoculated Medicago truncatula roots, containing
                     a mixture of young and old roots and nodules."
                     /note="Vector: Lambda Zap; Four-week old Rhizobium
BASE COUNT    182 a    135 c    167 g    177 t
ORIGIN
Query Match    38.5%;  Score 25.4;  DB 9;  Length 661;
Best Local Similarity   68.68%;  Pred. No. 1e+02;
Matches   35;  Conservative   0;  Mismatches   16;  Indels   0;  Gaps   0;

Qy    4 GATCTCGGATCCGTTCGACAAAACGGTAAAAAAGCCGTAGAITACCGTCCTA 54
         |||||||||||||| ||||||| |||||||||||||||||| |||||||||||
Db   67 GATTCTGTTCCGTTCGATAAAACCGATTGTAAAAACGATGTCTTATCGTCCTA 117

RESULT 11
BF642214
```

Right section

```
LOCUS       BF642214                680 bp    mRNA        linear    EST 19-DEC-2000
DEFINITION  NF067C04IN1F1033 Insect herbivory Medicago truncatula cDNA clone
            NF067C04IN 5', mRNA sequence.
ACCESSION   BF642214
VERSION     BF642214.1  GI:11906372
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Korth,K., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores
            ,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula insect herbivory library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Korth K
            Dept. of Plant Pathology
            University of Arkansas
            217 Plant Science Building, Fayetteville, AR 72701, USA
            Tel: 501 575 5191
            Fax: 501 575 7601
            Email: kkorth@comp.uark.edu
            Insert Length: 680  Std Error: 0.00
            Plate: 067  row: C  column: 04
            Seq primer: TCACACAGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1..680
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF067C04IN"
                     /clone_lib="Insect herbivory"
                     /tissue_type="local and systemic leaves"
                     /dev_stage="mature"
                     /note="Vector: Lambda Zap; Library was produced from fully
                     expanded M. truncatula leaves of plants fed upon by
                     Spodoptera exigua (beet armyworm) for 24 hours. Systemic
                     (undamaged leaves from injured plants) and wounded leaves
                     were harvested and pooled."
BASE COUNT    182 a    139 c    174 g    180 t    5 others
ORIGIN
Query Match    38.5%;  Score 25.4;  DB 10;  Length 680;
Best Local Similarity   68.68%;  Pred. No. 1e+02;
Matches   35;  Conservative   0;  Mismatches   16;  Indels   0;  Gaps   0;

Qy    4 GATCTCGGATCCGTTCGACAAAACGGTAAAAAAGCCGTAAAAAAACGGTAGAITACCGTCCTA 54
         |||||||||||||| |||||||||||||||| ||| ||||||||||||||| |||||||||||
Db   70 GATTCTGTTCCGTTCGATAAAACCGATTGTAAAAACGATGTCTTATCGTCCTA 120

RESULT 12
LOCUS       AW689706                688 bp    mRNA        linear    EST 15-JUN-2000
            AW689706
DEFINITION  NF023D11S1F1000 Developing stem Medicago truncatula cDNA clone
            NF023D11ST 5', mRNA sequence.
ACCESSION   AW689706
VERSION     AW689706.1  GI:7564442
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 688)
  AUTHORS   He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell
            ,C.J., Flores,H.R., Inman,J.T., Weller,J.W., May,G.D. and Dixon
            ,R.A.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula stem library
```

JOURNAL   Unpublished (2000)
COMMENT   Contact: Dixon RA
           Plant Biology Division
           The Samuel Roberts Noble Foundation
           2510 Sam Noble Parkway, Ardmore, OK 74402, USA
           Tel: 580 221 7302
           Fax: 580 221 7380
           Email: radixon@noble.org
           Insert Length: 688   Std Error: 0.00
           Plate: 023 Row: D Column: 11
           Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES
     source   Location/Qualifiers
               1..688
               /organism="Medicago truncatula"
               /db_xref="taxon:3880"
               /clone="NF023D11ST"
               /clone_lib="Developing stem"
               /tissue_type="stem"
               /dev_stage="Pooled developmental"
               /note="Vector: Lambda Zap; Contains a mixture of
               internodal stem segments"
BASE COUNT   181 a   139 c   176 g   189 t   3 others
ORIGIN

Qy   4 GATCTCGATCCGTTCGACAAAAACGGTAAAAAAGCCGTAGAATTACCGTCCTA 54
        |||||| ||||||||||||||| |||| ||| | | ||||||||| |||||||
Db 57 GATTCTGCTTCCGTTCGATAAAACCGATTCTAAAAACCATGTCTTTATCGTCCCTA 107

Query Match 48.5%;   Score 25.4;   DB 9;   Length 688;
Best Local Similarity 68.6%;   Pred. No. 1e+02;
Matches 35;   Conservative 0;   Mismatches 16;   Indels 0;   Gaps 0;

RESULT 13
LOCUS   BG645754   777 bp   mRNA   linear   EST 24-APR-2001
DEFINITION EST507473 KV3 Medicago truncatula cDNA clone pKV3-47H11 5' end,
            mRNA sequence.
ACCESSION   BG645754
VERSION   BG645754.1 GI:13780866
KEYWORDS   EST.
SOURCE   barrel medic.
  ORGANISM Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1 (bases 1 to 777)
  AUTHORS   VandenBosch,K., Endre,G., Hur,J., Heremand,P., Town,C.D., Bowman
            ,C.L., Craven,M.H., Cho,J. and Fraser,C.M.
  TITLE   ESTs from roots of Medicago truncatula 72 h after Rhizobium
            inoculation.
  JOURNAL   Unpublished (2001)
COMMENT   Contact: VandenBosch K
            Department of Biology
            Texas A&M University
            College Station, TX 77843-3258, USA
            Tel: 409 845 7707
            Fax: 409 845 2891
            Email: katev@mail.bio.tamu.edu
            M3941616 Tick sequence name: MTECG061TK More information is
            available at: www.medicago.org
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES
     source   Location/Qualifiers
               1..777
               /organism="Medicago truncatula"
               /cultivar="genotype A17"
               /db_xref="taxon:3880"
               /clone="pKV3-47H11"
               /clone_lib="KV3"
               /tissue_type="seedling roots"
               /dev_stage="3 days post-inoculation with Sinorhizobium
               meliloti."

Qy   4 GATCTCGATCCGTTCGACAAAAACGGTAAAAAAGCCGTAGAATTACCGTCCTA 54
        |||||| |||||||||||||| ||||| ||| | | |||||||||||||||||
Db 51 GATTCTGCTTCCGTTCGATAAAACCGATTCTAAAAACCATGTCTTTATCGTCCCTA 101

Query Match 48.5%;   Score 25.4;   DB 9;   Length 777;
Best Local Similarity 68.6%;   Pred. No. 1e+02;
Matches 35;   Conservative 0;   Mismatches 16;   Indels 0;   Gaps 0;

BASE COUNT   206 a   158 c   196 g   217 t
ORIGIN

               /lab_host="E. coli strain XL1BlR"
               /note="Vector: pBluescript SK -; Site_1: EcoRI; Site_2:
               XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
               was directionally ligated into the Unizap XR vector from
               Stratagene and packaged using Gigapack III Gold packaging
               extracts. Plasmids containing cDNA inserts were excised
               from the recombinant lambda-Zap phage using Ex-assist
               helper phage and propagated in XL0lR cells."

RESULT 14
LOCUS   AI260450   459 bp   mRNA   linear   EST 19-APR-2001
DEFINITION LP04321.5prime LP Drosophila melanogaster cDNA clone LP04321 5prime similar to
            Drosophila melanogaster cDNA clone LP04321 5prime, mRNA
            sequence.
ACCESSION   AI260450
VERSION   AI260450.1 GI:3867975
KEYWORDS   EST.
SOURCE   fruit fly.
  ORGANISM Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 459)
  AUTHORS   Harvey,D., Hrokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE   BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 43 Row: B Column: 9
            High quality sequence stop: 422.
FEATURES
     source   Location/Qualifiers
               1..459
               /organism="Drosophila melanogaster"
               /db_xref="taxon:7227"
               /clone="LP04321"
               /clone_lib="LP Drosophila melanogaster larval-early pupal
               pOT2"
               /sex="male and female"
               /dev_stage="larvae-pupae"
               /lab_host="DH5-alpha"
               /note="Organ: whole body; Vector: pOT2; Site_1: EcoRI;
               Site_2: XhoI; Sized fractionated cDNAs were directly
               ligated into pOT2. Plasmid cDNA library."

Qy   4 GATCTCGATCCGTTCGACAAAAACGGTAAAAAAGCCGTAGAATTACCGTCCTA 54
        |||||| |||||||||||||| || ||| | | |||||||||||||||||
Db 51 GATTCTGCTTCCGTTCGATAAAACCGATTCTAAAAACCATGTCTTTATCGTCCCTA 101

Query Match 48.5%;   Score 25.4;   DB 10;   Length 459;
Best Local Similarity 68.6%;   Pred. No. 1.1e+02;
Matches 35;   Conservative 0;   Mismatches 16;   Indels 0;   Gaps 0;

BASE COUNT   97 a   127 c   147 g   88 t
ORIGIN

Qy   8 TGGATCGTTCGACAAAAACGGTAAAAAGCCGTAGATTACCGTCTCTTACGAAGTTC 61
        ||||||| ||||| ||| |||| ||||| ||| || ||| ||||| ||| ||||
Db 66 TGGATCCAATGGACCAGCAGAGCGGAAACAGCCGGACAATACCGCCTGGACACATTC 119

```
RESULT 15
AA698485

LOCUS       AA698485             521 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  HL04775.5prime HL Drosophila melanogaster head Bluescript
            Drosophila melanogaster cDNA clone HL04775 5prime similar to
            sequence.
ACCESSION   AA698485
VERSION     AA698485.1   GI:2701414
KEYWORDS    EST
SOURCE      fruit fly.
ORGANISM    Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 521)
AUTHORS     Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
TITLE       BDGP/HHMI Drosophila EST Project
JOURNAL     Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            HDGP
            Lawrence Berkeley National Lab
            one Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Plate: 47 row: C column: 3
            High quality sequence stop: 378.
FEATURES             Location/Qualifiers
     source          1..521
                     /organism="Drosophila melanogaster"
                     /db_xref="taxon:7227"
                     /clone="HL04775"
                     /clone_lib="HL Drosophila melanogaster head BlueScript"
                     /sex="male and female"
                     /dev_stage="adult"
                     /lab_host="SOLR"
                     /note="Organ: head--brain & sensory organ; Vector:
                     BlueScript SK; Site_1: EcoRI; Site_2: XhoI; Constructed
                     using Stratagene ZAP-cDNA Synthesis kit. Oligo dT-primed
                     and directionally cloned at EcoRI and XhoI in BlueScript
                     SK(+/-)"
BASE COUNT      111 a    134 c    153 g    123 t
ORIGIN
```

Query Match   38.2%; Score 25.2; DB 9; Length 521;
Best Local Similarity   66.7%; Pred. No. 1.2e+02;
Matches   36; Conservative   0; Mismatches 18; Indels   0; Gaps   0;

Qy   8 TCGATCCGTTCGACAAAACGGTAAAAAAGCGGTAGATTACCGTCCTACGGAATTC   61
            ||||||| | |||||| ||||| ||||||| |||| |||| |||| |||
Db   419 TCGATCCAATGGACGACGAAGCGAAAACAGGCCGGAGAATACCGCTGGACACATTC 472

Search completed: November 5, 2002, 08:09:53
Job time : 808.027 secs
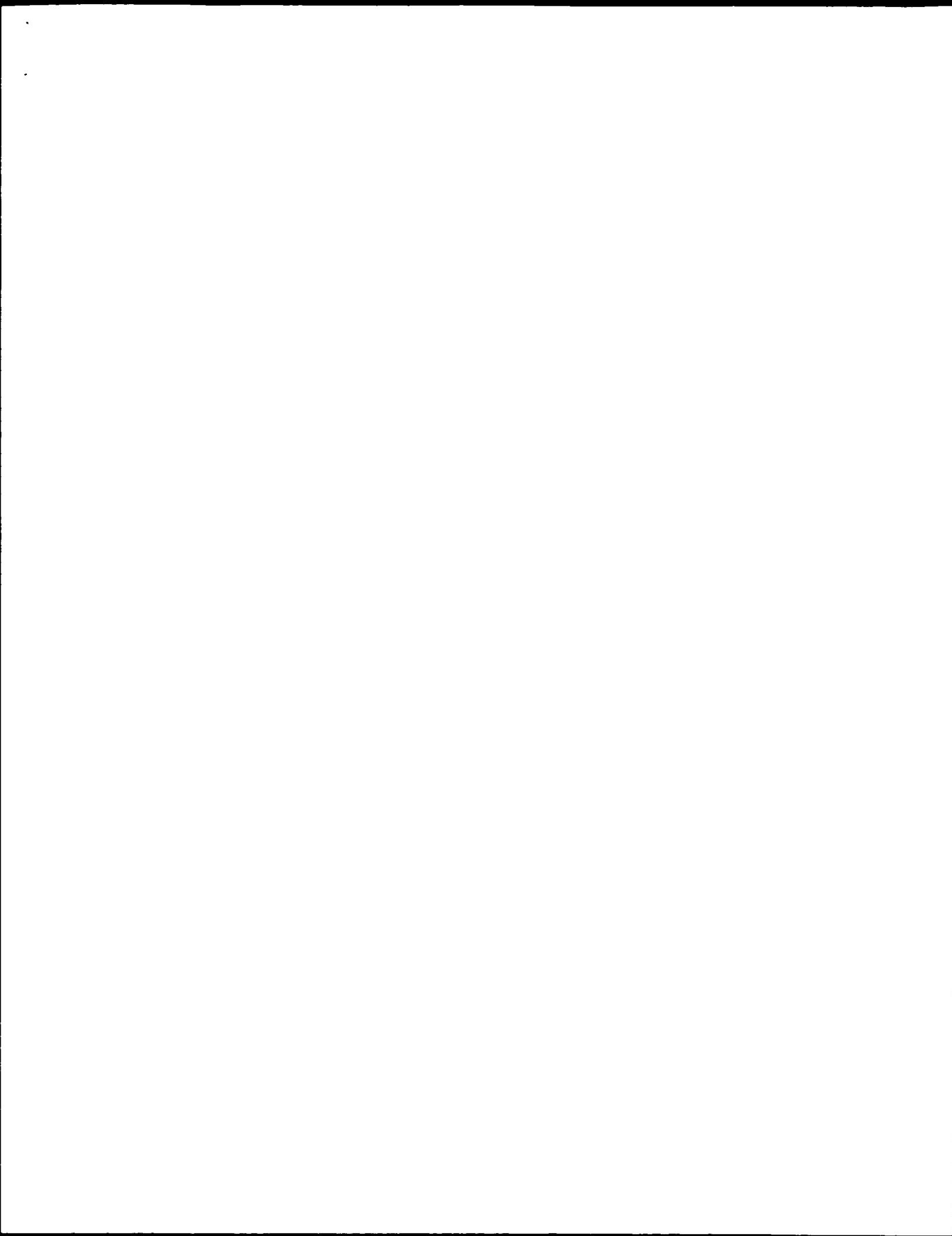
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 05:27:49 ; Search time 300.568 seconds
(without alignments)
4595.145 Million cell updates/sec

Title: US-09-766-113-2
Perfect score: 66
Sequence: 1 tgaaatctggaatccagttcga.............ccgtcctacgaattcaactg 66

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 1046326829 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: qb_ba:*
2: qb_htg:*
3: qb_in:*
4: qb_om:*
5: qb_ov:*
6: qb_pat:*
7: qb_ph:*
8: qb_pl:*
9: qb_pr:*
10: qb_ro:*
11: qb_sts:*
12: qb_sy:*
13: qb_un:*
14: qb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_or:*
21: em_om:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

SUMMARIES

Result
No. Score Query Match length DB ID Description
-------- ----- ----- ------------- -- -- -----------

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

	Score						
1	66	66	100.0	160666	6	AX203094	AX203094 Sequence
2	66	66	100.0	162069	2	AX203116	AX203116 Sequence
3	48.4	73.3	66.6	168247	3	AX203095	AX203095 Sequence
4	37.8	57.3	66.6	195937	9	AX203113	AX203113 Sequence
5	37.2	56.4	66.6	26	6	AX203093	AX203093 Sequence
6	37.2	56.4	66.6	47200	2	AX207072	AX207072 Sequence
7	37.2	56.4	66.6	56268	1	AC013076	AC013076 Drosophil
8	37.2	56.4	66.6	171620	1	AC074316	AC074316 Staphyloc
9	36.2	54.8	66.6	291829	1	AC010070	AC010070 Drosophil
10	36	54.5	66.6	343590	1	AP003359	AP003359 Staphyloc
11	35.6	53.9	66.6	65	6	AX203106	AX203106 Sequence
12	35.6	53.9	66.6	4226	3	DROMDR49	DROMDR49 Drosophil
13	35.6	53.9	66.6	4229	3	AY058791	AY058791 Drosophil
14	34.6	52.4	66.6	39097	2	AC014852	AC014852 Drosophil
15	26.6	40.3	66.6	163052	3	AC007472	AC007472 Drosophil
16	26.6	40.3	66.6	265387	3	AE003820	AE003820 Drosophil
17	26.6	40.3	66.6	972	1	STHISK	STHISK
18	26.6	40.3	66.6	4378	1	STYCARAHA	STYCARAHA
19	26	39.4	66.6	20621	1	AE008883	AE008883 Salmonell
20	26	39.4	66.6	46699	9	AL136135	AL136135 Human DNA
21	26	39.4	66.6	49114	2	AC024336	AC024336 Homo sapi
22	26	39.4	66.6	59320	3	AC024746	AC024746 Caenorhab
23	26	39.4	66.6	69462	2	AC101335	AC101335 Mus muscu
24	26	39.4	66.6	96086	2	AF233324	AF233324 Salmonell
25	25.2	38.2	65	99653	2	AC103038	AC103038 Rattus no
26	25.2	38.2	66.6	102945	2	AL389884	AL389884 Homo sapi
27	25.2	38.2	66.6	105713	2	AL392020	AL392020 Homo sapi
28	25.2	38.2	66.6	153180	2	AC024916	AC024916 Homo sapi
29	25.2	38.2	66.6	180664	2	AC006706	AC006706 Caenorhab
30	25.2	38.2	66.6	7329	8	AF035007	AF035007 Kluyverom
31	25	37.9	66	84419	2	AC006245	AC006245 Drosophil
32	26.6	40.3	53.9			AX203108	
33	26.6	40.3	53.9			AC073025	
34	26	39.4	52.4			AC096762	
35	26	39.4	47200			AX207072	
36	37.2	56.4				AX203109	
37	37.2	56.4				AX203110	
38	37.2	56.4				AX203115	
39	37.2	56.4				AX203111	
40	36.2	54.8				AX203101	
41	35.6	53.9				AX203107	
42	35.6	53.9				AX203112	
43	35.6	53.9				AX203114	
44	34.6	52.4				AC068959	
45	26.6	40.3				AC108928	

ALIGNMENTS

RESULT 1
AX203094

LOCUS   AX203094   66 bp   DNA   linear   PAT   30-AUG-2001
DEFINITION   Sequence 2 from patent WO0153502.
ACCESSION   AX203094
VERSION   AX203094.1   GI:15392453
KEYWORDS   .
SOURCE   synthetic construct.
  ORGANISM   synthetic construct
             artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 2 26-JUL-2001;
             PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="random oligonucleotide"
BASE COUNT   21 a   14 c   16 g   15 t
ORIGIN

RESULT 2

```
                 Query Match          100.0%;  Score 66;  DB 6;  Length 66;
AX204116         Best Local Similarity  100.0%;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
LOCUS        AX204116                   66 bp    DNA    linear    PAT 30-AUG-2001
DEFINITION   Sequence 24 from Patent WO0153502.
ACCESSION    AX204116
VERSION      AX204116.1  GI:15392477
KEYWORDS     .
SOURCE       synthetic construct.
  ORGANISM   synthetic construct
             artificial sequence.
REFERENCE    1  (bases 1 to 66)
  AUTHORS    Bruce,W.B. and Niu,X.
  TITLE      Novel root-preferred promoter elements and methods of use
  JOURNAL    Patent: WO 0153502-A 24 26-JUL-2001;
             PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT      21 a      14 c     16 g     15 t
ORIGIN
Qy   1 TGAGATCTGGATCCGTTCGACAAAAACGGGTAGATTACCGTCCTACGAATT  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGAGATCTGGATCCGTTCGACAAAAACGGGTAGATTACCGTCCTACGAATT  60

Qy  61 CAGCTG  66
         ||||||
Db  61 CAGCTG  66
```

RESULT 3

```
                 Query Match          100.0%;  Score 66;  DB 6;  Length 66;
AX203095         Best Local Similarity  100.0%;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;
LOCUS        AX203095                   66 bp    DNA    linear    PAT 30-AUG-2001
DEFINITION   Sequence 3 from Patent WO0153502.
ACCESSION    AX203095
VERSION      AX203095.1  GI:15392454
KEYWORDS     .
SOURCE       synthetic construct.
  ORGANISM   synthetic construct
             artificial sequence.
REFERENCE    1  (bases 1 to 66)
  AUTHORS    Bruce,W.B. and Niu,X.
  TITLE      Novel root-preferred promoter elements and methods of use
  JOURNAL    Patent: WO 0153502-A 3 26-JUL-2001;
             PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="random oligonucleotide"
BASE COUNT      22 a      14 c     15 g     15 t
ORIGIN
Qy   1 TGAGATCTGGATCCGTTCGACAAAAACGGGTAGATTACCGTCCTACGAATT  60
         ||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGAGATCTGGATCCGTTCGACAAAAACGGGTAGATTACCGTCCTACGAATT  60

Qy  61 CAGCTG  66
         ||||||
Db  61 CAGCTG  66
```

RESULT 4

```
                 Query Match           74.3%;  Score 48.4;  DB 6;  Length 66;
AX203113         Best Local Similarity  83.3%;  Pred. No. 2.1e-07;  Mismatches 11;  Indels 0;  Gaps 0;
LOCUS        AX203113                   66 bp    DNA    linear    PAT 30-AUG-2001
DEFINITION   Sequence 21 from Patent WO0153502.
ACCESSION    AX203113
VERSION      AX203113.1  GI:15392474
KEYWORDS     .
SOURCE       synthetic construct.
  ORGANISM   synthetic construct
             artificial sequence.
REFERENCE    1  (bases 1 to 66)
  AUTHORS    Bruce,W.B. and Niu,X.
  TITLE      Novel root-preferred promoter elements and methods of use
  JOURNAL    Patent: WO 0153502-A 21 26-JUL-2001;
             PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT      20 a      12 c     20 g     13 t     1 others
ORIGIN
Qy   1 TGAGATCTGGATCCGTTCGACAAAAACGGGTAGATTACCGTCCTACGAATT  60
         |||||||||||||||||||||||| |   |  |  ||||||||||||||||
Db   1 TGAGATCTGGATCCGTTCGACAAAACTAAAACTAAAACGTAACTGACCTGCCTCCTACGAATT  60

Qy  61 CAGCTG  66
         ||||||
Db  61 CAGCTG  66
```

RESULT 5

```
                 Query Match           57.3%;  Score 37.8;  DB 6;  Length 66;
AX203093         Best Local Similarity  73.8%;  Pred. No. 0.0018;  Mismatches 17;  Indels 0;  Gaps 0;
LOCUS        AX203093                   66 bp    DNA    linear    PAT 30-AUG-2001
DEFINITION   Sequence 1 from Patent WO0153502.
ACCESSION    AX203093
VERSION      AX203093.1  GI:15392451
KEYWORDS     .
SOURCE       synthetic construct.
  ORGANISM   synthetic construct
             artificial sequence.
REFERENCE    1  (bases 1 to 66)
  AUTHORS    Bruce,W.B. and Niu,X.
  TITLE      Novel root-preferred promoter elements and methods of use
  JOURNAL    Patent: WO 0153502-A 1 26-JUL-2001;
             PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="random oligonucleotide"
BASE COUNT      19 a      12 c     21 g     14 t
ORIGIN
Qy   2 GAGATCTGGATCCGTTCGACAAAAACGGCTAAAAAGCGGGTAGATTACCGTCCTACGAATTC  61
        ||||||||||||||||||||||||  ||    |   | |   ||||||||||||||||||||
Db   2 GAGATCTGGATCCGATCCCGTTCGGGGGAAGGGCAAGGTGAAAGCTAAAAATTACCGTCCTACGAATTC  61

Qy  62 AGCTG  66
        |||||
Db  62 AGCTG  66
```

RESULT 6

Query Match   56.4%;   Score 37.2;   DB 6;   Length 66;
Best Local Similarity   72.7%;   Pred. No. 0.003;
Matches   48;   Conservative   0;   Mismatches   18;   Indels   0;   Gaps   0;

```
Qy      1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAACGGTAGATTACCGTCCTACGAATT 60
          |||||||||||||||||   ||   ||  | |  || |||||||||||||||||||||
Db      1 TGAGATCTGGATCCGTTCGGGGAAGGAGGTGAAACCAAGAATTACCGTCCTACGAATT 60

Qy     61 CAGCTG 66
          ||||||
Db     61 CAGCTG 66
```

LOCUS   AX203109   66 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION   Sequence 17 from Patent WO0153502.
ACCESSION   AX203109
VERSION   AX203109.1   GI:15392470
KEYWORDS   .
SOURCE   synthetic construct.
   ORGANISM   synthetic construct
     artificial sequence.
REFERENCE   1 (bases 1 to 66)
   AUTHORS   Bruce,W.B. and Niu,X.
   TITLE   Novel root-preferred promoter elements and methods of use
   JOURNAL   Patent: WO 0153502-A 17 26-JUL-2001;
     PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
   source   1..66
     /organism="synthetic construct"
     /db_xref="taxon:32630"
     /note="synthetic sequences flanking a random
     oligonucleotide"
BASE COUNT   19 a   12 c   21 g   14 t
ORIGIN

RESULT 7

Query Match   56.4%;   Score 37.2;   DB 6;   Length 66;
Best Local Similarity   72.7%;   Pred. No. 0.003;
Matches   48;   Conservative   0;   Mismatches   18;   Indels   0;   Gaps   0;

```
Qy      1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAACGGTAGATTACCGTCCTACGAATT 60
          |||||||||||||||||   ||        || |  || |||||||||||||||||||
Db      1 TGAGATCTGGATCCGTTCGGGGGAAGGGAAGGTGAAAGCAAGAATTACCGTCCTACGAATT 60

Qy     61 CAGCTG 66
          ||||||
Db     61 CAGCTG 66
```

LOCUS   AX203110   66 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION   Sequence 18 from Patent WO0153502.
ACCESSION   AX203110
VERSION   AX203110.1   GI:15392471
KEYWORDS   .
SOURCE   synthetic construct.
   ORGANISM   synthetic construct
     artificial sequence.
REFERENCE   1 (bases 1 to 66)
   AUTHORS   Bruce,W.B. and Niu,X.
   TITLE   Novel root-preferred promoter elements and methods of use
   JOURNAL   Patent: WO 0153502-A 18 26-JUL-2001;
     PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
   source   1..66
     /organism="synthetic construct"
     /db_xref="taxon:32630"
     /note="synthetic sequences flanking a random
     oligonucleotide"
BASE COUNT   19 a   12 c   21 g   14 t
ORIGIN

RESULT 8

Query Match   56.4%;   Score 37.2;   DB 6;   Length 66;
Best Local Similarity   72.7%;   Pred. No. 0.003;
Matches   48;   Conservative   0;   Mismatches   18;   Indels   0;   Gaps   0;

```
Qy      1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAACGGTAGATTACCGTCCTACGAATT 60
          |||||||||||||||||   ||     | |  || |||||||||||||||||||||||
Db      1 TGAGATCTGGATCCGTTCGGGGAAGGGAAGGTGAAAGCAAGAATTACCGTCCTACGAATT 60

Qy     61 CAGCTG 66
          ||||||
Db     61 CAGCTG 66
```

LOCUS   AX203115   66 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION   Sequence 23 from Patent WO0153502.
ACCESSION   AX203115
VERSION   AX203115.1   GI:15392476
KEYWORDS   .
SOURCE   synthetic construct.
   ORGANISM   synthetic construct
     artificial sequence.
REFERENCE   1 (bases 1 to 66)
   AUTHORS   Bruce,W.B. and Niu,X.
   TITLE   Novel root-preferred promoter elements and methods of use
   JOURNAL   Patent: WO 0153502-A 23 26-JUL-2001;
     PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
   source   1..66
     /organism="synthetic construct"
     /db_xref="taxon:32630"
     /note="synthetic sequences flanking a random
     oligonucleotide"
BASE COUNT   19 a   12 c   22 g   13 t
ORIGIN

RESULT 9

Query Match   56.4%;   Score 37.2;   DB 6;   Length 66;
Best Local Similarity   72.7%;   Pred. No. 0.003;
Matches   48;   Conservative   0;   Mismatches   18;   Indels   0;   Gaps   0;

```
Qy      1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAACGGTAGATTACCGTCCTACGAATT 60
          |||||||||||||||||   ||        || |  || |||||||||||||||||||
Db      1 TGAGATCTGGATCCGTTCGGGAGAAGGGAAGGTGAAAGCAGGAAAATACCGTCCTACGAATT 60

Qy     61 CAGCTG 66
          ||||||
Db     61 CAGCTG 66
```

LOCUS   AX203111   66 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION   Sequence 19 from Patent WO0153502.
ACCESSION   AX203111
VERSION   AX203111.1   GI:15392472
KEYWORDS   .
SOURCE   synthetic construct.
   ORGANISM   synthetic construct
     artificial sequence.
REFERENCE   1 (bases 1 to 66)
   AUTHORS   Bruce,W.B. and Niu,X.
   TITLE   Novel root-preferred promoter elements and methods of use
   JOURNAL   Patent: WO 0153502-A 19 26-JUL-2001;
     PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
   source   1..66
     /organism="synthetic construct"
     /db_xref="taxon:32630"
     /note="synthetic sequences flanking a random
     oligonucleotide"
BASE COUNT   19 a   11 c   21 g   14 t   1 others

ORIGIN

RESULT 10
AX204101
    Query Match   54.8%; Score 36.2; DB 6; Length 66;
    Best Local Similarity   71.2%; Pred. No. 0.007;
    Matches 47; Conservative   0; Mismatches 19; Indels   0; Gaps   0;

Qy   1 TGAGATCTGGATCTCCGTTCGACAAAAYGGTAAAAAAGCGGTAGATTACCGTCCTACGGAATT 60
              ||||||||||| |||||| |||||| || || || ||||||||||||||||||||||||
Db   1 TGAGATCTGGATCNGTTCGGGGAACGGAAGGTGAAAGCAAGAATTACCGTCCTACGGAATT 60

Qy   61 CAGCTG 66
              ||||||
Db   61 CAGCTG 66

RESULT 11
AX204107
LOCUS   AX204107   66 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION Sequence 15 from Patent WO0153502.
ACCESSION   AX204107
VERSION   AX204107.1 GI:15392468
KEYWORDS   .
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 15 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT   20 a   12 c   21 g   13 t

RESULT 12
AX203108
LOCUS   AX203108   66 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION Sequence 16 from Patent WO0153502.
ACCESSION   AX203108
VERSION   AX203108.1 GI:15392469
KEYWORDS   .
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 16 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT   19 a   11 c   21 g   15 t

RESULT 13
AX203112
LOCUS   AX203112   66 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION Sequence 20 from Patent WO0153502.
ACCESSION   AX203112
VERSION   AX203112.1 GI:15392473
KEYWORDS   .
SOURCE   synthetic construct.
  ORGANISM synthetic construct
            artificial sequence.
REFERENCE   1 (bases 1 to 66)
  AUTHORS   Bruce,W.B. and Niu,X.
  TITLE   Novel root-preferred promoter elements and methods of use
  JOURNAL   Patent: WO 0153502-A 20 26-JUL-2001;
            PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"

ORIGIN

    Query Match   53.9%; Score 35.6; DB 6; Length 66;
    Best Local Similarity   71.2%; Pred. No. 0.012;
    Matches 47; Conservative   0; Mismatches 19; Indels   0; Gaps   0;

Qy   1 TGAGATCTGGATCTCCGTTCGACAAAACGGTAAAAAAGCGGTAGATTACCGTCCTACGGAATT 60
              ||||||||||| |||||| |||||| | || || ||||||||||||||||||||||||
Db   1 TGAGATCTGGATTCGTTCGGGGAACGGAAGGTGAAAGCAAGAATTACCGTCCTACGGAATT 60

Qy   61 CAGCTG 66
              ||||||
Db   61 CAGCTG 66

BASE COUNT   19 a   11 c   21 g   15 t
ORIGIN

RESULT 13

Query Match   51.9%;   Score 35.6;   DB 6;   Length 66;
Best Local Similarity   71.2%;   Pred. No. 0.012;
Matches 47;   Conservative   0;   Mismatches 19;   Indels 0;   Gaps 0;

Qy   1 TGAGATCTGGATCCGTTCGGTTCGACAAAAACGGTAAAAAACGGTAGATTACCGTCCTACGAATT 60
           ||||||||||||||||||| | |||||||| |||||| ||||||||||||||| ||||||||||
Db   1 TGAGATCTGGATCCGTTCGGTTCGGGGAAGGAAGCAAGCCAGAATTACTGTCCTACGAATT 60

Qy   61 CAGCTG 66
           ||||||
Db   61 CAGCTG 66

RESULT 14

AX203114
LOCUS   AX203114   66 bp   DNA   linear   PAT 30-AUG-2001
DEFINITION   Sequence 22 from Patent WO0153502.
ACCESSION   AX203114
VERSION   AX203114.1   GI:15392475
KEYWORDS
SOURCE   synthetic construct.
ORGANISM   synthetic construct
              artificial sequence.
REFERENCE   1 (bases 1 to 66)
AUTHORS   Bruce,W.B. and Niu,X.
TITLE   Novel root-preferred promoter elements and methods of use
JOURNAL   Patent: WO 0153502-A 22 26-JUL-2001;
              PIONEER HI-BRED INTERNATIONAL, INC. (US)
FEATURES   Location/Qualifiers
     source   1..66
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="synthetic sequences flanking a random
                     oligonucleotide"
BASE COUNT   19 a   12 c   21 g   13 t   1 others
ORIGIN

Query Match   52.4%;   Score 34.6;   DB 6;   Length 66;
Best Local Similarity   70.8%;   Pred. No. 0.027;
Matches 46;   Conservative   0;   Mismatches 19;   Indels 0;   Gaps 0;

Qy   2 GAGATCTGGATCCGTTCGACAAAAACGGTAAAAAAGCGGTAGATTACCGTCCTACGAATTC 61
           ||||||||||||||||||| | ||||| || | || ||||||||||||||||||||||||
Db   2 GAGATCTGGATCCGGGAAGGGAAGCTGAAAGTAAGAATTACCGTCTCCTACGAATCC 61

RESULT 15

AC068959
LOCUS   AC068959   160666 bp   DNA   linear   HTG 04-JUN-2000
DEFINITION   Homo sapiens clone RP11-12010, WORKING DRAFT SEQUENCE, 11 unordered
              pieces.
ACCESSION   AC068959
VERSION   AC068959.2   GI:8247876
KEYWORDS   HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   human.
ORGANISM   Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 160666)
AUTHORS   Hirten,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE   Homo sapiens, clone RP11-12010
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 160666)
AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
              Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
              Boguslavkiy,L., Bhadwin,J., Brown,A., Burkett,G.,

Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., LaRocque,R., Lamazares,M.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierce,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teslaye,S., Theodore,J., Tirrell,A., Travers,M., Triglia,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

COMMENT   On Jun 4, 2000 this sequence version replaced gi:7798795.

              Center: Whitehead Institute/MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
     ----
              -------------- Project Information --------------
              Center Project Name: L7444
              Center Clone Name: 12_0_10

              -------------- Summary Statistics --------------
              Sequencing Vector: M13; M77815; 100% of reads
              Chemistry: Dye-terminator Big Dye; 100% of reads
              Assembly program: Phrap; version 0.960731
              Consensus quality: 153117 bases at least Q40
              Consensus quality: 157272 bases at least Q30
              Consensus quality: 158589 bases at least Q20
              Insert size: 163000; agarose-fp
              Insert size: 159666; sum-of-contigs
              Quality coverage: 4.5 in Q20 bases; agarose-fp
              Quality coverage: 4.6 in Q20 bases; sum-of-contigs
     ----
              NOTE: This is a 'working draft' sequence. It currently
     *   consists of 11 contigs. The true order of the pieces
     *   is not known and their order in this sequence record is
     *   arbitrary. Gaps between the contigs are represented as
     *   runs of N, but the exact sizes of the gaps are unknown.
     *   This record will be updated with the finished sequence
     *   as soon as it is available and the accession number will
     *   be preserved.
     *
     *   1   1072: contig of 1072 bp in length
     *   1073   1172: gap of   100 bp
     *   1173   2585: contig of 1413 bp in length
     *   2586   2685: gap of   100 bp
     *   2686   5841: contig of 3156 bp in length
     *   5842   5941: gap of   100 bp
     *   5942 16755: contig of 10814 bp in length
     *   16756 16855: gap of   100 bp
     *   16856 31901: contig of 15046 bp in length
     *   31902 32001: gap of   100 bp
     *   32002 50422: contig of 18421 bp in length
     *   50423 50522: gap of   100 bp
     *   50523 67507: contig of 16985 bp in length
     *   67508 67607: gap of   100 bp
     *   67608 87476: contig of 19869 bp in length
     *   87477 87576: gap of   100 bp
     *   87577 107074: contig of 19498 bp in length
     *   107075 107174: gap of   100 bp
     *   107175 132166: contig of 24992 bp in length
     *   132167 132266: gap of   100 bp

```
FEATURES             Location/Qualifiers
     *    142267    160666: contig of 28400 bp in length.
     source         1..160666
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone_lib="RPll-120l0"
                    /clone="RPCI-11 Human Male BAC"
     misc_feature   1..1072
                    /note="assembly_fragment"
     misc_feature   1171..2585
                    /note="assembly_fragment"
     misc_feature   2686..5841
                    /note="assembly_fragment"
     misc_feature   5942..16755
                    /note="assembly_fragment"
     misc_feature   16856..31901
                    /note="assembly_fragment"
                    clone_end:SP6
                    vector_side:right
     misc_feature   32002..50422
                    /note="assembly_fragment"
     misc_feature   50524..67507
                    /note="assembly_fragment"
     misc_feature   67608..87476
                    /note="assembly_fragment"
     misc_feature   87577..107074
                    /note="assembly_fragment"
                    clone_end:T7
                    vector_side:right
     misc_feature   107175..132166
                    /note="assembly_fragment"
     misc_feature   142267..160666
                    /note="assembly_fragment"
BASE COUNT      51792 a   29510 c   29548 g   48787 t   1009 others
ORIGIN
Query Match       40.4%;   Score 26.6;   DB 2;   Length 160666;
Best Local Similarity   64.1%;   Pred. No. 45;
Matches   41;   Conservative   0;   Mismatches   24;   Indels   0;   Gaps   0;

Qy          1 TGACATCTGGATCCGTTCGACTAAAAACGGTAAAAAAGCGGTAGATTACCGTCCTACGAATT 60
              ||  ||| |||||| ||| |||| | ||||   |   ||| | ||| ||| || |||| |||
Db       8414 TGTGATTTTGTTTCGTTCCATCAAGTTTTACATAAACGTAAGCTTATCTTCCCACTAATA 8354

Qy         61 CAGGT 65
              ||| |
Db       8454 CAGAT 8449

Search completed: November 5, 2002, 06:26:38
Job time : 458.568 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 05:25:34 ; Search time 95.9421 Seconds
                  (without alignments)
                  1181.091 Million cell updates/sec

Title: us-09-766-113-2
Perfect score: 66
Sequence: 1 tagaatctgaatccgttcga.........ccgtcctacgatcaagctg 66

Scoring table: IDENTITY NUC
                  Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

Database :   N_Geneseq_032802:*

```
 1: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1980.DAT:*
 2: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1981.DAT:*
 3: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1982.DAT:*
 4: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1983.DAT:*
 5: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1984.DAT:*
 6: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1985.DAT:*
 7: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1986.DAT:*
 8: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1987.DAT:*
 9: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcqdata/geneseq/geneseqn-embl/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match length	DB	ID	Description
1	66	100.0	66	22	AAS08433	A plant root-prefe
2	66	100.0	66	22	AAS08449	A plant root-prefe
3	48.4	73.3	66	22	AAS08434	A plant root-prefe
4	37.8	57.3	66	22	AAS08450	A plant root-prefe
5	37.2	56.4	66	22	AAS08432	A plant root-prefe
6	37.2	56.4	66	22	AAS08452	A plant root-prefe
7	36.2	54.8	66	22	AAS08448	A plant root-prefe
8	36	54.5	66	22	AAS08440	A random oligonucl
9	35.6	53.9	66	22	AAS08446	A plant root-prefe
10	35.6	53.9	66	22	AAS08447	A plant root-prefe
11	35.6	53.9	66	22	AAS08449	A plant root-prefe
12	34.6	52.4	66	22	AAS08451	A plant root-prefe
13	34.6	52.4	66	22	AAH42733	A promoter element
14	26	39.4	26	18	AAV74611	Staphylococcus aur
15	25.2	38.2	65	22	AAS08445	A plant root-prefe
16	25.2	38.2	65	22	AAH53570	DNA encoding novel
17	25.2	38.2	66	23	ABL06835	Drosophila melanog
18	25.2	38.2	66	23	ABL06834	Drosophila melanog
19	24.4	37.0	66	22	AAS73285	S. epidermidis ope
20	24.4	37.0	66	22	AAS81874	A plant root-prefe
21	24.4	37.0	66	22	AAH54503	S. epidermidis gen
22	24.4	37.0	66	21	AAZ55699	DNA encoding yeast
23	24.4	37.0	66	19	AAV43879	Mycobacterial meth
24	24.2	36.7	66	23	ABL18043	Drosophila melanog
25	24.2	36.7	66	23	ABL18042	Drosophila melanog
26	24.2	36.7	66	22	AAI99683	A plant root-prefe
27	24	36.4	66	23	AAS08437	Mycobacterium tube
28	23.8	36.1	66	23	ABL07964	Drosophila melanog
29	23.6	35.8	66	23	AAS70677	DNA encoding novel
30	23.6	35.8	66	23	AAS79513	DNA encoding novel
31	23.6	35.8	66	23	AAS70040	DNA encoding novel
32	23.6	35.8	66	24	AHL32184	Human immune syste
33	23.4	35.5	66	21	AAC94671	Cat flea hindgut a
34	23.4	35.5	66	22	AAS08439	A plant root-prefe
35	23.2	35.2	66	21	ABL09554	Drosophila melanog
36	23.2	35.2	66	21	AAA81474	N. meningitidis pa
37	23.2	35.2	66	21	AAF21611	Neisseria meningit
38	23.2	35.2	66	21	AAA81489	N. meningitidis pa
39	23	34.8	66	24	AAS61380	Human gene regulat
40	23	34.8	66	23	AAS61380	Drosophila melanog
41	22.4	33.9	66	21	AAA01453	Human colon cancer
42	22.4	33.9	66	23	AAA81849	DNA encoding novel
43	22.4	33.9	66	23	AAS69854	DNA encoding novel
44	22.4	33.9	66	22	AAF60852	Pseudomonas sp exp
45	22.4	33.9	66	20	AAX04212	pAP-222 insert con

ALIGNMENTS

RESULT 1
AAS08433
AAS08433;

```
ID   AAS08433 standard; DNA; 66 BP.
XX
AC   AAS08433;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 14.
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE14; random oligonucleotide library;
OS   Synthetic.
XX
FH   Key             location/Qualifiers
FH
FT   misc_feature    1..18
FT                   /*tag= a
FT                   /label= "5'_flanking_sequence"
FT   misc_feature    19..48
FT                   /*tag= b
FT                   /note= "Randomised sequence"
FT   misc_feature    49..66
FT                   /*tag= c
FT                   /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
```

XX
PF 19-JAN-2001; 2001W0-US02011.
XX
PR 21-JAN-2000; 2000US-0177473.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Bruce WB,
PI Niu X;
XX
DR WPI; 2001-442261/47.
XX
PT Producing tissue-preferred promoter elements constructs for regulating
PT expression of nucleotide sequences in a plant comprises identifying and
PT isolating tissue-preferred promoter elements -
XX
PS Claim 5; Fig 1; 45pp; English.
XX
CC The sequence represents a plant root-preferred promoter element, RPE,
CC isolated from a random oligonucleotide library (ROL). The invention
CC concerns a method of identifying and isolating tissue-preferred promoter
CC elements comprising the use of a mixture of random oligonucleotides,
CC flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC specific plant nuclear proteins, isolating the complexes and PCR
CC amplifying the bound oligonucleotide. The method is used for isolating
CC tissue-specific promoters from plants, including but not limited to
CC root-specific promoters or root-preferred promoter elements (RPE). The
CC RPEs are useful in the genetic manipulation of a plant when operably
CC linked to a nucleotide sequence whose expression is to be controlled to
CC achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC temperature, salinity, pesticide and herbicide resistance) and biotic
CC stress (disease resistance, resistance to attack by fungi, bacteria,
CC viruses, insects and nematodes).
XX
SQ Sequence 66 BP; 21 A; 14 C; 16 G; 15 T; 0 other;

Query Match   100.0%;   Score 66;   DB 22;   Length 66;
Best Local Similarity   100.0%;   Pred. No. 2.7e-16;
Matches 66;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGTAGAGATTACCGTCCGTCCTACGGAATT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGTAGAGATTACCGTCCGTCCTACGGAATT 60

Qy 61 CAGCTG 66
        ||||||
Db 61 CAGCTG 66

RESULT 2
AAS08454
ID AAS08454 standard; DNA; 66 BP.
XX
AC AAS08454;
XX
DT 26-SEP-2001 (first entry)
XX
DE A plant root-preferred promoter element (RPE), RPE 21.
XX
KW Root preferred promoter element; RPE; abiotic stress; drought;
KW salinity; pesticide resistance; herbicide resistance; biotic stress;
KW disease resistance; fungal disease; bacterial disease; viral disease;
KW insect attack; nematode attack; RPE21; random oligonucleotide library;
KW ROL; ss.
XX
OS Synthetic.
XX
FH Key   Location/Qualifiers
FH
FT misc_feature   1..18
FT   /*tag= a
FT   /label= "5'_flanking_sequence"
FT misc_feature   19..48
FT   /*tag= b
FT   /note= "Randomised sequence"

FT misc_feature   49..66
FT   /*tag= c
FT   /label= "3'_flanking_sequence"
XX
PN WO200153502-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001W0-US02011.
XX
PR 21-JAN-2000; 2000US-0177473.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Bruce WB,
PI Niu X;
XX
DR WPI; 2001-442261/47.
XX
PT Producing tissue-preferred promoter elements constructs for regulating
PT expression of nucleotide sequences in a plant comprises identifying and
PT isolating tissue-preferred promoter elements -
XX
PS Example 1; Fig 1; 45pp; English.
XX
CC The sequence represents a plant root-preferred promoter element, RPE,
CC isolated from a random oligonucleotide library (ROL). The invention
CC concerns a method of identifying and isolating tissue-preferred promoter
CC elements comprising the use of a mixture of random oligonucleotides,
CC flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC specific plant nuclear proteins, isolating the complexes and PCR
CC amplifying the bound oligonucleotide. The method is used for isolating
CC tissue specific promoters from plants, including but not limited to
CC root-specific promoters or root-preferred promoter elements (RPE). The
CC RPEs are useful in the genetic manipulation of a plant when operably
CC linked to a nucleotide sequence whose expression is to be controlled to
CC achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC temperature, salinity, pesticide and herbicide resistance) and biotic
CC stress (disease resistance, resistance to attack by fungi, bacteria,
CC viruses, insects and nematodes).
XX
SQ Sequence 66 BP; 21 A; 14 C; 16 G; 15 T; 0 other;

Query Match   100.0%;   Score 66;   DB 22;   Length 66;
Best Local Similarity   100.0%;   Pred. No. 2.7e-16;
Matches 66;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGTAGAGATTACCGTCCGTCCTACGGAATT 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGTAGAGATTACCGTCCGTCCTACGGAATT 60

Qy 61 CAGCTG 66
        ||||||
Db 61 CAGCTG 66

RESULT 4
AAS08434
ID AAS08434 standard; DNA; 66 BP.
XX
AC AAS08434;
XX
DT 26-SEP-2001 (first entry)
XX
DE A plant root-preferred promoter element (RPE), RPE 19.
XX
KW Root-preferred promoter element; RPE; abiotic stress; drought;
KW salinity; pesticide resistance; herbicide resistance; biotic stress;
KW disease resistance; fungal disease; bacterial disease; viral disease;
KW insect attack; nematode attack; RPE19; random oligonucleotide library;
KW ROL; ss.
XX
OS Synthetic.
XX

```
FH   Key             Location/Qualifiers
FH
FT   misc_feature    1..18
FT                   /*tag= a
FT                   /label= "5'_flanking_sequence"
FT   misc_feature    19..48
FT                   /*tag= b
FT                   /note= "Randomised sequence"
FT   misc_feature    49..66
FT                   /*tag= c
FT                   /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
DR   WPI; 2001-442261/47.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB,   Niu X;
XX
PT   claim 5; Fig 1; 45pp; English.
XX
PS
XX
The sequence represents a plant root-preferred promoter element, RPE,
isolated from a random oligonucleotide library (ROL). The invention
concerns a method of identifying and isolating tissue-preferred promoter
elements comprising the use of a mixture of random oligonucleotides,
flanked by 5' and 3' flanking sequences, which hybridise to tissue-
specific plant nuclear proteins, isolating the complexes and PCR
amplifying the bound oligonucleotide. The method is used for isolating
tissue-specific promoters from plants, including but not limited to
root-specific promoters or root-preferred promoter elements (RPE). The
RPEs are useful in the genetic manipulation of a plant when operably
linked to a nucleotide sequence whose expression is to be controlled to
achieve a desired phenotypic effect, e.g. abiotic stress (drought,
temperature, salinity, pesticide and herbicide resistance) and biotic
stress (disease resistance, resistance to attack by fungi, bacteria,
viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 22 A; 14 C; 15 G; 15 T; 0 other;

Query Match             73.3%;  Score 48.4;  DB 22;  Length 66;
Best Local Similarity   83.3%;
     Matches  55;  Conservative   0;  Mismatches  11;  Indels   0;  Gaps   0;

Qy           1 TCGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGGTAGAATTACCGTCCTACGAATT 60
               |||||||||||||||||||||||||||||||||  |||||| | ||||||||||||||||||
Db           1 TCGAGATCTGGATCCGTTCGACAAAACGGTAAAACTAAAGGTAACTGACGTCCTACGAATT 60

Qy          61 CAGCTG 66
               ||||||
Db          61 CAGCTG 66
```

```
RESULT 4
AAS08450
ID   AAS08450 standard; DNA; 66 BP.
XX
AC   AAS08450;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE); RPE 33.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
```

```
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE33; random oligonucleotide library;
KW   ROL; ss.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FH
FT   misc_feature    1..18
FT                   /*tag= a
FT                   /label= "5'_flanking_sequence"
FT   misc_feature    19..48
FT                   /*tag= b
FT                   /note= "Randomised sequence"
FT   misc_feature    49..66
FT                   /*tag= c
FT                   /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB,   Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Example 1; Fig 1; 45pp; English.
XX
PS
XX
The sequence represents a plant root-preferred promoter element, RPE,
isolated from a random oligonucleotide library (ROL). The invention
concerns a method of identifying and isolating tissue-preferred promoter
elements comprising the use of a mixture of random oligonucleotides,
flanked by 5' and 3' flanking sequences, which hybridise to tissue-
specific plant nuclear proteins, isolating the complexes and PCR
amplifying the bound oligonucleotide. The method is used for isolating
tissue-specific promoters from plants, including but not limited to
root-specific promoters or root-preferred promoter elements (RPE). The
RPEs are useful in the genetic manipulation of a plant when operably
linked to a nucleotide sequence whose expression is to be controlled to
achieve a desired phenotypic effect, e.g. abiotic stress (drought,
temperature, salinity, pesticide and herbicide resistance) and biotic
stress (disease resistance, resistance to attack by fungi, bacteria,
viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 20 A; 12 C; 20 G; 13 T; 1 other;

Query Match             57.3%;  Score 37.8;  DB 22;  Length 66;
Best Local Similarity   73.8%;
     Matches  48;  Conservative   0;  Mismatches  17;  Indels   0;  Gaps   0;

Qy           2 GAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGGTAGAATTACCGTCCTACGAATTC 61
               || |||||||||||||||| |  |||| ||  ||||||| |||||||||||||||||||||
Db           2 GACATCTGGATCCGTTCGGGGAAGGCTGAAAGCAAAAATTACCCTCCTACGAATTC 61

Qy          62 AGCTG 66
               |||||
Db          62 AGCTG 66
```

```
RESULT 5
AAS08432
ID   AAS08432 standard; DNA; 66 BP.
XX
```

AC   AAS08452;
XX
DT   26-SEP-2001 (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 15.
XX
KW   Root preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE15; random oligonucleotide library;
KW   ROL; ss.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT
FT   misc_feature   1..18
FT   /*tag= a
FT   /label= "5'_flanking_sequence"
FT   misc_feature   19..48
FT   /*tag= b
FT   /note= "Randomised sequence"
FT   misc_feature   49..66
FT   /*tag= c
FT   /label= "3'_flanking_sequence"
XX
PN   (PION) PIONEER HI-BRED INT INC.
XX
PA   (PION) PIONEER HI-BRED INT INC.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PD   26-JUL-2001.
XX
PD   19-JAN-2001; 2001WO-US02011.
XX
DR   WO200153502-A2.
XX
PN   Bruce WB, Niu X;
XX
PI   WPI; 2001-442261/47.
XX
PS   Claim 5; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 19 A; 12 C; 21 G; 14 T; 0 other;

Query Match   56.4%; Score 37.2; DB 22; Length 66;
Best Local Similarity   72.7%; Pred. No. 4.6e-05;
  Matches 48; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy   1 TGA-ANCTGGAGATCCNTGCAACAAANCGGTAAAAAAGCGGTAGATTACGTCCCTACGAATT 60
                 || |||||||||||||||| | |||||||| | | || ||||||||||||||||||||
Db   1 TGAGATCTGGAGATCCNTCGGCGAAGGGAAGGTGAAGCTGAAAGCAAGCAAATTACCGTCCCTACGAATT 60

Qy   61 CAGCTG 66
                 ||||||
Db   CAGCTG 66

Db   61 CAGCTG 66

RESULT 6
AAS08452
ID   AAS08452 standard; DNA; 66 BP.
XX
AC   AAS08452;
XX
DT   26-SEP-2001 (first entry)
XX
DE   A plant root-preferred promoter element; RPE; abiotic stress; drought;
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE74; random oligonucleotide library;
KW   ROL; ss.
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT
FT   misc_feature   1..18
FT   /*tag= a
FT   /label= "5'_flanking_sequence"
FT   misc_feature   19..48
FT   /*tag= b
FT   /note= "Randomised sequence"
FT   misc_feature   49..66
FT   /*tag= c
FT   /label= "3'_flanking_sequence"
XX
PN   (PION) PIONEER HI-BRED INT INC.
XX
PA   (PION) PIONEER HI-BRED INT INC.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PD   26-JUL-2001.
XX
PD   19-JAN-2001; 2001WO-US02011.
XX
DR   WO200153502-A2.
XX
PN   Bruce WB, Niu X;
XX
PI   WPI; 2001-442261/47.
XX
PS   Example 1; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 19 A; 12 C; 22 G; 14 T; 0 other;

Query Match   56.4%; Score 37.2; DB 22; Length 66;
Best Local Similarity   72.7%; Pred. No. 4.6e-05;
  Matches 48; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

```
Qy    1 TGAGATCTGGATCGTTCGGACAAAACGGTAAAAAACGGGTAGATTACCGTCCTACGAATT 60
        |||||||||||||||   ||   |  |||||| |||||||||||||||||
Db    1 TGAGATCTGGACCGTTCGGACGAAGGGCAAAGGTGAAGGCAGGAAATACCGTCCTACGAATT 60

Qy   61 CAGCTG 66
        ||||||
Db   61 CAGCTG 66

RESULT 7
AAS08448
ID   AAS08448 standard; DNA; 66 BP.
XX
AC   AAS08448;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 89.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; RPE89; random oligonucleotide library;
KW   ROL; ss.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FH
FT   misc_feature    1..18
FT                   /tag=    a
FT                   /label=  "5' flanking_sequence"
FT   misc_feature    19..48
FT                   /tag=    b
FT                   /note=   "Randomised sequence"
FT   misc_feature    49..66
FT                   /tag=    c
FT                   /label=  "3' flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Fig 1; 45pp; English.
XX
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating root-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
```

```
CC   viruses, insects and nematodes).
XX
SQ   Sequence 66 BP; 19 A; 11 C; 21 G; 14 T; 1 other;

     Query Match      54.8%;  Score 36.2;  DB 22;  Length 66;
     Best Local Similarity  71.2%;  Pred. No. 0.00011;
     Matches  47;  Conservative  0;  Mismatches  19;  Indels  0;  Gaps  0;

Qy    1 TGAGATCTGGATCGTTCGGACAAAACGGTAAAAAACGGGTAGATTACCGTCCTACGAATT 60
        ||||||||||||  ||||||  |||  |  ||  |  |||||||||||||||||
Db    1 TGAGATCTGGATCGTCGGGGAAGGGGAAGGGGAAGGGAAAAGTAAAGAATTACCGTCCTACGAATT 60

Qy   61 CAGCTG 66
        ||||||
Db   61 CAGCTG 66

RESULT 8
AAS08440
ID   AAS08440 standard; DNA; 66 BP.
XX
AC   AAS08440;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A random oligonucleotide library, ROL, sequence n19813.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; bacterial disease; viral disease;
KW   insect attack; nematode attack; random oligonucleotide library;
KW   ROL; n19813; ss.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FH
FT   misc_feature    1..18
FT                   /tag=    a
FT                   /label=  "5' flanking_sequence"
FT   misc_feature    19..48
FT                   /tag=    b
FT                   /note=   "Randomised sequence"
FT   misc_feature    49..66
FT                   /tag=    c
FT                   /label=  "3' flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PT   Producing tissue-preferred promoter elements constructs for regulating
PT   expression of nucleotide sequences in a plant comprises identifying and
PT   isolating tissue-preferred promoter elements -
XX
PS   Example 1; Page 2l; 45pp; English.
XX
CC   The sequence represents a random oligonucleotide library (ROL) construct
CC   used to isolate tissue-specific promoter elements. The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
```

CC tissue-specific promoters from plants, including but not limited to
CC root-specific promoters or root-preferred promoter elements (RPE). The
CC RPEs are useful in the genetic manipulation of a plant when operably
CC linked to a nucleotide sequence whose expression is to be controlled to
CC achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC temperature, salinity, pesticide and herbicide resistance) and biotic
CC stress (disease resistance, resistance to attack by fungi, bacteria,
CC viruses, insects and nematodes).
XX
SQ Sequence 66 BP; 7 A; 9 C; 9 G; 11 T; 30 other;

Query Match   54.5%; Score 36; DB 22; Length 66;
Best Local Similarity   54.5%; Pred. No. 0.00014;
Matches 36; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

Qy 1 TGAGATCTGGATCCGTTCGACYAAAAYCGGTAAAAAACRCGGTAGCATTACCCGTCCTACGAATT 60
       ||||||||||||||||||||| |||||||||||||||
Db 1 TGAGATCTGGCATCCGTTCTCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNGTCCTACGAATT 60

Qy 61 CAGCTG 66
       ||||||
Db 61 CAGCTG 66

RESULT 9
AAS08446
ID AAS08446 standard; DNA; 66 BP.
XX
AC AAS08446;
XX
DT 26-SEP-2001 (first entry)
XX
DE A plant root-preferred promoter element (RPE), RPE 18.
XX
KW Root-preferred promoter element; RPE; abiotic stress; drought;
KW salinity; pesticide resistance; herbicide resistance; biotic stress;
KW disease resistance; fungal disease; bacterial disease; viral disease;
KW insect attack; nematode attack; RPE18; random oligonucleotide library;
KW ROL; ss.
OS Synthetic.
XX
FH Key   Location/Qualifiers
FH
FT misc_feature   1..18
FT   /*tag= a
FT   /label= "5'_flanking_sequence"
FT misc_feature   19..48
FT   /*tag= b
FT   /note= "Randomised sequence"
FT misc_feature   49..66
FT   /*tag= c
FT   /label= "3'_flanking_sequence"
XX
PN WO200153502-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US02011.
XX
PR 21-JAN-2000; 2000US-0177473.
XX
PA (PION-) PIONEER HI BRED INT INC.
XX
PI Bruce WB, Niu X;
XX
DR WPI; 2001-442261/47.
XX
PT Producing tissue-preferred promoter elements constructs for regulating
PT expression of nucleotide sequences in a plant comprises identifying and
PT isolating tissue-preferred promoter elements.
XX
PS Example 1: Fig 1; 45pp; English.

The sequence represents a plant root-preferred promoter element, RPE,
isolated from a random oligonucleotide library (ROL). The invention
concerns a method of identifying and isolating tissue-preferred promoter
elements comprising the use of a mixture of random oligonucleotides,
flanked by 5' and 3' flanking sequences, which hybridise to tissue-
specific plant nuclear proteins, isolating the complexes and PCR
amplifying the bound oligonucleotide. The method is used for isolating
tissue-specific promoters from plants, including but not limited to
root-specific promoters or root-preferred promoter elements (RPE). The
RPEs are useful in the genetic manipulation of a plant when operably
linked to a nucleotide sequence whose expression is to be controlled to
achieve a desired phenotypic effect, e.g. abiotic stress (drought,
temperature, salinity, pesticide and herbicide resistance) and biotic
stress (disease resistance, resistance to attack by fungi, bacteria,
viruses, insects and nematodes).
XX
SQ Sequence 66 BP; 20 A; 12 C; 21 G; 13 T; 0 other;

Query Match   45.9%; Score 45.6; DB 22; Length 66;
Best Local Similarity   71.2%; Pred. No. 0.00019;
Matches 47; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 1 TGAGATCTGGATCCGTTCGACAAAAACCGTAAAAAAGRGGTAGCATTACCCGTCCTACGAATT 60
       ||||||||||| |||||||| | || | ||||||||||||||||||||||
Db 1 TGAGATCTGGGAACCGTTCGGGCGAAGGAAGGTCAAAACAAAGAATTACCCGTCCTACGAATT 60

Qy 61 CAGCTG 66
       ||||||
Db 61 CAGCTG 66

RESULT 10
AAS08447
ID AAS08447 standard; DNA; 66 BP.
XX
AC AAS08447;
XX
DT 26-SEP-2001 (first entry)
XX
DE A plant root-preferred promoter element (RPE), RPE 32.
XX
KW Root-preferred promoter element; RPE; abiotic stress; drought;
KW salinity; pesticide resistance; herbicide resistance; biotic stress;
KW disease resistance; fungal disease; bacterial disease; viral disease;
KW insect attack; nematode attack; RPE32; random oligonucleotide library;
KW ROL; ss.
OS Synthetic.
XX
FH Key   Location/Qualifiers
FH
FT misc_feature   1..18
FT   /*tag= a
FT   /label= "5'_flanking_sequence"
FT misc_feature   19..48
FT   /*tag= b
FT   /note= "Randomised sequence"
FT misc_feature   49..66
FT   /*tag= c
FT   /label= "3'_flanking_sequence"
XX
PN WO200153502-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US02011.
XX
PR 21-JAN-2000; 2000US-0177473.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Bruce WB, Niu X;
XX
DR WPI; 2001-442261/47.

XX
PT Producing tissue-preferred promoter elements constructs for regulating
PT expression of nucleotide sequences in a plant comprises identifying and
PT isolating tissue-preferred promoter elements -
XX
PS Example 1; Fig 1; 45pp; English.
XX
CC The sequence represents a plant root-preferred promoter element, RPE,
CC isolated from a random oligonucleotide library (ROL). The invention
CC concerns a method of identifying and isolating tissue-preferred promoter
CC RPEs are useful in the genetic manipulation of a plant when operably
CC elements comprising the use of a mixture of random oligonucleotides,
CC linked to a nucleotide sequence whose expression is to be controlled to
CC flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC specific plant nuclear proteins, isolating the complexes and PCR
CC temperature, salinity, pesticide and herbicide resistance) and biotic
CC amplifying the bound oligonucleotide. The method is used for isolating
CC stress (disease resistance, resistance to attack by fungi, bacteria,
CC tissue-specific promoters from plants, including but not limited to
CC viruses, insects and nematodes).
XX
SQ Sequence 66 BP; 19 A; 11 C; 21 G; 15 T; 0 other;

Query Match   53.9%; Score 35.6; DB 22; Length 66;
Best Local Similarity 71.2%; Pred. No. 0.00019;
Matches 47; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy   1 TCGAGATCTGGATCCGTTCGGACAAAAACGGTAAAAAAGCGGTAGATTACGTCGTACGAATT 60
        |||||||||||||| |||||| || ||||||||||||||||||||||||
Db   1 TGAGATCTGGATTCGTTCGGGGGAAGGGAAGGTGAAAACGCAAGAATTACGTCCTACGAATT 60

Qy 61 CAGCTG 66
        ||||||
Db 61 CAGCTG 66

RESULT 11
AAS08449
ID AAS08449 standard; DNA; 66 BP.
XX
AC AAS08449;
XX
DT 26-SEP-2001 (first entry)
XX
DE A plant root-preferred promoter element (RPE), RPE 22.
XX
KW Root-preferred promoter element; RPE; abiotic stress; drought;
KW salinity; pesticide resistance; herbicide resistance; biotic stress;
KW disease resistance; fungal disease; bacterial disease; viral disease;
KW insect attack; nematode attack; RPE22; random oligonucleotide library;
KW ROL; ss.
OS Synthetic.
XX
FH Key   Location/Qualifiers
FH
FT misc_feature   1..18
FT   /*tag=   a
FT   /label= "5'_flanking_sequence"
FT misc_feature   19..48
FT   /*tag=   b
FT   /note=   "Randomised sequence"
FT misc_feature   49..66
FT   /*tag=   c
FT   /label= "3'_flanking_sequence"
XX
PN WO200153502-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US02011.
XX

PR 21-JAN-2000; 2000US-0177473.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Bruce WB, Niu X;
XX
DR WPI; 2001-442261/47.
XX
PS Example 1; Fig 1; 45pp; English.
XX
PT Producing tissue-preferred promoter elements constructs for regulating
PT expression of nucleotide sequences in a plant comprises identifying and
PT isolating tissue-preferred promoter elements -
XX
CC The sequence represents a plant root-preferred promoter element, RPE,
CC isolated from a random oligonucleotide library (ROL). The invention
CC concerns a method of identifying and isolating tissue-preferred promoter
CC elements comprising the use of a mixture of random oligonucleotides,
CC flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC specific plant nuclear proteins, isolating the complexes and PCR
CC amplifying the bound oligonucleotide. The method is used for isolating
CC tissue-specific promoters from plants, including but not limited to
CC root-specific promoters or root-preferred promoter elements (RPE). The
CC RPEs are useful in the genetic manipulation of a plant when operably
CC linked to a nucleotide sequence whose expression is to be controlled to
CC achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC temperature, salinity, pesticide and herbicide resistance) and biotic
CC stress (disease resistance, resistance to attack by fungi, bacteria,
CC viruses, insects and nematodes).
XX
SQ Sequence 66 BP; 19 A; 11 C; 21 G; 15 T; 0 other;

Query Match   53.9%; Score 35.6; DB 22; Length 66;
Best Local Similarity 71.2%; Pred. No. 0.00019;
Matches 47; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy   1 TCGAGATCTGGATCCGTTCGACAAAAACGGTAAAAAAGCGGTAGATTACGTCGTACGAATT 60
        |||||||||||||| ||||| || |||||||||||||| |||||||||||
Db   1 TGAGATCTGGATCCGTTCGGGGGAAGGGAAGGTGAAAACGCAAGAATTACGTCCTACGAATT 60

Qy 61 CAGCTG 66
        ||||||
Db 61 CAGCTG 66

RESULT 12
AAS08451
ID AAS08451 standard; DNA; 66 BP.
XX
AC AAS08451;
XX
DT 26-SEP-2001 (first entry)
XX
DE A plant root-preferred promoter element (RPE), RPE 71.
XX
KW Root-preferred promoter element; RPE; abiotic stress; drought;
KW salinity; pesticide resistance; herbicide resistance; biotic stress;
KW disease resistance; fungal disease; bacterial disease; viral disease;
KW insect attack; nematode attack; RPE71; random oligonucleotide library;
KW ROL; ss.
OS Synthetic.
XX
FH Key   Location/Qualifiers
FH
FT misc_feature   1..18
FT   /*tag=   a
FT   /label= "5'_flanking_sequence"
FT misc_feature   19..48
FT   /*tag=   b
FT   /note=   "Randomised sequence"
FT misc_feature   49..66
FT   /*tag=   c
FT   /label= "3'_flanking_sequence"

XX
PN W02001543502-A2.
PD 26.JUL.2001.
PO
XX
PF
XX
PR
XX 19-JAN-2001; 2001WO-US020011.
XX
PR 21.JAN.2000; 2000US-0177473.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Bruce WB, Niu X;
XX
DR WPI; 2001-442261/47.
XX
PS Example 1; Fig 1; 45pp; English.
XX
PT Producing tissue-preferred promoter elements constructs for regulating
PT expression of nucleotide sequences in a plant comprises identifying and
PT isolating tissue-preferred promoter elements.
XX
CC The sequence represents a plant root-preferred promoter element, RPE,
CC isolated from a random oligonucleotide library (ROL). The invention
CC concerns a method of identifying and isolating tissue-preferred promoter
CC elements comprising the use of a mixture of random oligonucleotides,
CC flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC specific plant nuclear proteins, isolating the complexes and PCR
CC amplifying the bound oligonucleotide. The method is used for isolating
CC tissue-specific promoters from plants, including but not limited to
CC root-specific promoters or root-preferred promoter elements (RPE). The
CC RPEs are useful in the genetic manipulation of a plant when operably
CC linked to a nucleotide sequence whose expression is to be controlled to
CC achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC temperature, salinity, pesticide and herbicide resistance) and biotic
CC stress (disease resistance, resistance to attack by fungi, bacteria,
CC viruses, insects and nematodes).
XX
SQ Sequence 66 BP; 19 A; 12 C; 21 G; 13 T; 1 other;

Query Match 52.4%; Score 74.6; DB 22; Length 66;
Host Local Similarity 70.8%; Pred. No. 0.00048;
Matches 46; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 2 GACATGTTGGATVCGTTYCGACAAAACXGGTAAAAANCXGGTAGATTACCGTCCTACGSAATTC 61
      |||||||||||||||||| || ||| || ||| |||||||||||||||||||||||
Db 2 GACATTCGAATCCGTTCGGGGCGAAGCGGTAAAACTTAAGAATTAACCGTCCTACGAAATCC 61

Qy 62 AGRTTG 66
      |||||
Db 62 AGRTTG 66

RESULT 14
AAH42733
ID AAH42733 standard; DNA; 26 BP.
XX
AC AAH42733;
XX
DT 01-OCT-2001 (first entry)
XX
DE A promoter element or transcription binding site.
XX
KW Promoter element; transcription binding site; plant promoter; SMPER;
KW synthetic multimeric promoter element region; gene expression;
KW insect resistance; herbicide resistance; ss.
XX
OS Synthetic.
XX
PN W02001543476-A2.
XX
PD 26-JUL-2001.
XX
PF
XX
PF 19-JAN-2001; 2001WO-US020024.

XX
PR 21-JAN-2000; 2000US-0177473.
XX
PA (PION-) PIONEER HI-BRED INT INC.
XX
PI Bruce WB, Niu X;
XX
WPI; 2001-476118/51.
XX
DR New plant promoters with synthetic multimeric promoter element regions,
DR useful in plant molecular biology, particularly in regulating gene
XX expression in plants to increase resistance against insects or
XX herbicides
PT
XX
PS Example 1; Fig 1; 67pp; English.
XX
CC AAH42709-72 represent promoter elements or transcription binding sites.
CC They are used to to construct synthetic multimeric promoter element
CC regions (SMPERs). The specification describes plant promoters which
CC comprise SMPERs. The plant promoters are useful in plant molecular
CC biology, particularly in regulating gene expression in plants. The
CC promoters are especially useful for transforming plants or plant cells,
CC e.g. to increase resistance against insects or herbicides.
XX
SQ Sequence 26 BP; 12 A; 4 C; 6 G; 4 T; 0 other;

Query Match 39.4%; Score 26; DB 22; Length 26;
Host Local Similarity 100.0%; Pred. No. 0.87;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 23 AAACGGTAAAAAANCGGTAGATTACC 48
       ||||||||||||||||||||||||||
Db 1 AAACGGTAAAAAANCGGTAGATTACC 26

RESULT 14
AAV74611
ID AAV74611 standard; DNA; 3181 BP.
XX
AC AAV74611;
XX
DT 16-MAR-1999 (first entry)
XX
DE Staphylococcus aureus contig SEQ ID #300.
XX
KW Computer readable medium; vaccine; S.aureus infection; immunodetection;
KW cellulitis; eyelid infection; food poisoning; osteomyelitis; therapy;
KW skin infection; surgical wound infection; scalded skin syndrome;
KW toxic shock syndrome; ds.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FH misc_feature
FT misc_feature 181..240
FT /*taq=a
FT /note= "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
FT misc_feature 1981..2040
FT /*taq=b
FT /note- "these bases represent a line of missing text in
FT the sequence listing in the specification. They
FT are included to maintain the nucleotide numbering
FT given in the specification for this DNA sequence"
XX
PN EP786519-A2.
XX
PD 30-JUL-1997.
XX
PF 07-JAN-1997; 97EP-0100117.
XX
PR 05-JAN-1996; 96US-0009861.

```
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Barash SC,  Choi GH,  Dillon PJ,  Fannon MR,  Kunsch CA;
PI   Rosen CA;
XX
DR   WPI; 1997-374922/35.
XX
PS   Claim 1: Page 1128-1130; 3271pp; English.
XX
CC   This sequence represents one ot 5191 Staphylococcus aureus DNA sequences
CC   of the invention. The DNA sequences are recorded on a computer readable
CC   medium, preferably selected from a floppy or hard disk, random access
CC   memory (RAM), read-only memory (ROM) or CD-ROM. Homology searches using
CC   the S.aureus DNA sequences allows putative functions to be assigned so
CC   that protein-encoding or regulatory regions of commercial, therapeutic or
CC   industrial importance can be obtained. Specifically, sequences which are
CC   likely to encode antigens have been identified and these polypeptides can
CC   be used in a vaccine composition against S.aureus infection. The
CC   polypeptides can also be used in a kit for the immunodetection of
CC   S.aureus in a sample. S.aureus is implicated in numerous human diseases,
CC   including cellulitis, eyelid infections, food poisoning, osteomyelitis,
CC   skin and surgical wound infections, scalded skin syndrome, toxic shock
CC   syndrome, etc. Organisms transformed with the DNA sequences can be used
CC   for recombinant production of the polypeptides. The new DNA sequences
CC   (and their fragments) are useful as primers or probes for isolating
CC   homologues of any of the S.aureus DNA sequences contained on the
CC   computer readable medium.
XX
SQ   Sequence 3181 BP; 1005 A; 549 C; 483 G; 1022 T; 122 other;

Query Match   39.4%;  Score 26;  DB 18;  Length 3181;
Best Local Similarity  62.1%;  Pred. No. 2.6;
Matches  41;  Conservative  0;  Mismatches  25;  Indels  0;  Gaps  0;

Qy    1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGGTAGATTACCGTCCTACGGAATT 60
          ||| || ||||| ||  || | ||||||||| ||| |||||||||| || |||||||||||
Db 1134 TAATATATGGACCAGTTATTCTTAATTGGTTTACAGAGCATTAGANTATGTACCTAAAAAAG 1193

Qy   61 CAGCTG 66
         ||||||
Db 1194 CAGCTG 1199
```

```
RESULT 15
AAS08445
ID   AAS08445 standard; DNA; 65 BP.
XX
AC   AAS08445;
XX
DT   26-SEP-2001  (first entry)
XX
DE   A plant root-preferred promoter element (RPE), RPE 5.
XX
KW   Root-preferred promoter element; RPE; abiotic stress; drought;
KW   salinity; pesticide resistance; herbicide resistance; biotic stress;
KW   disease resistance; fungal disease; viral disease;
KW   insect attack; nematode attack; RPE5; random oligonucleotide library;
KW   ROL 5s.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FH
FT   misc_feature    1..18
FT                   /*tag= a;
FT                   /label= "5'_flanking_sequence"
FT   misc_feature    19..47
FT                   /*tag= b
FT                   /note= "Randomised sequence"
```

```
FT   misc_feature    48..65
FT                   /*tag= c
FT                   /label= "3'_flanking_sequence"
XX
PN   WO200153502-A2.
XX
PD   26-JUL-2001.
XX
PF   19-JAN-2001; 2001WO-US02011.
XX
PR   21-JAN-2000; 2000US-0177473.
XX
PA   (PION-) PIONEER HI-BRED INT INC.
XX
PI   Bruce WB, Niu X;
XX
DR   WPI; 2001-442261/47.
XX
PS   Example 1: Fig 1; 45pp; English.
XX
CC   Producing tissue-preferred promoter elements constructs for regulating
CC   expression of nucleotide sequences in a plant comprises identifying and
CC   isolating tissue-preferred promoter elements -
CC   The sequence represents a plant root-preferred promoter element, RPE,
CC   isolated from a random oligonucleotide library (ROL). The invention
CC   concerns a method of identifying and isolating tissue-preferred promoter
CC   elements comprising the use of a mixture of random oligonucleotides,
CC   flanked by 5' and 3' flanking sequences, which hybridise to tissue-
CC   specific plant nuclear proteins, isolating the complexes and PCR
CC   amplifying the bound oligonucleotide. The method is used for isolating
CC   tissue-specific promoters from plants, including but not limited to
CC   root-specific promoters or root-preferred promoter elements (RPE). The
CC   RPEs are useful in the genetic manipulation of a plant when operably
CC   linked to a nucleotide sequence whose expression is to be controlled to
CC   achieve a desired phenotypic effect, e.g. abiotic stress (drought,
CC   temperature, salinity, pesticide and herbicide resistance) and biotic
CC   stress (disease resistance, resistance to attack by fungi, bacteria,
CC   viruses, insects and nematodes).
XX
SQ   Sequence 65 BP; 18 A; 14 C; 17 G; 16 T; 0 other;

Query Match   38.2%;  Score 25.2;  DB 22;  Length 65;
Best Local Similarity  71.2%;  Pred. No. 2.2;
Matches  47;  Conservative  0;  Mismatches  18;  Indels  1;  Gaps  1;

Qy    1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGGTAGATTACCGTCCTACGGAATT 60
          |||||||||||||||||||||||||  ||  || | ||||||||| || ||||||||||||
Db    1 TCAGATCTGGATCCGTTCGAGCAGTAAAAGTAAGAAAAGGCCCGTT-TCGTCCTACGGAATT 59

Qy   61 CAGCTG 66
         ||||||
Db   60 CAGCTG 65

Search completed: November  5, 2002, 05:41:46
Job time : 100.942 secs
```
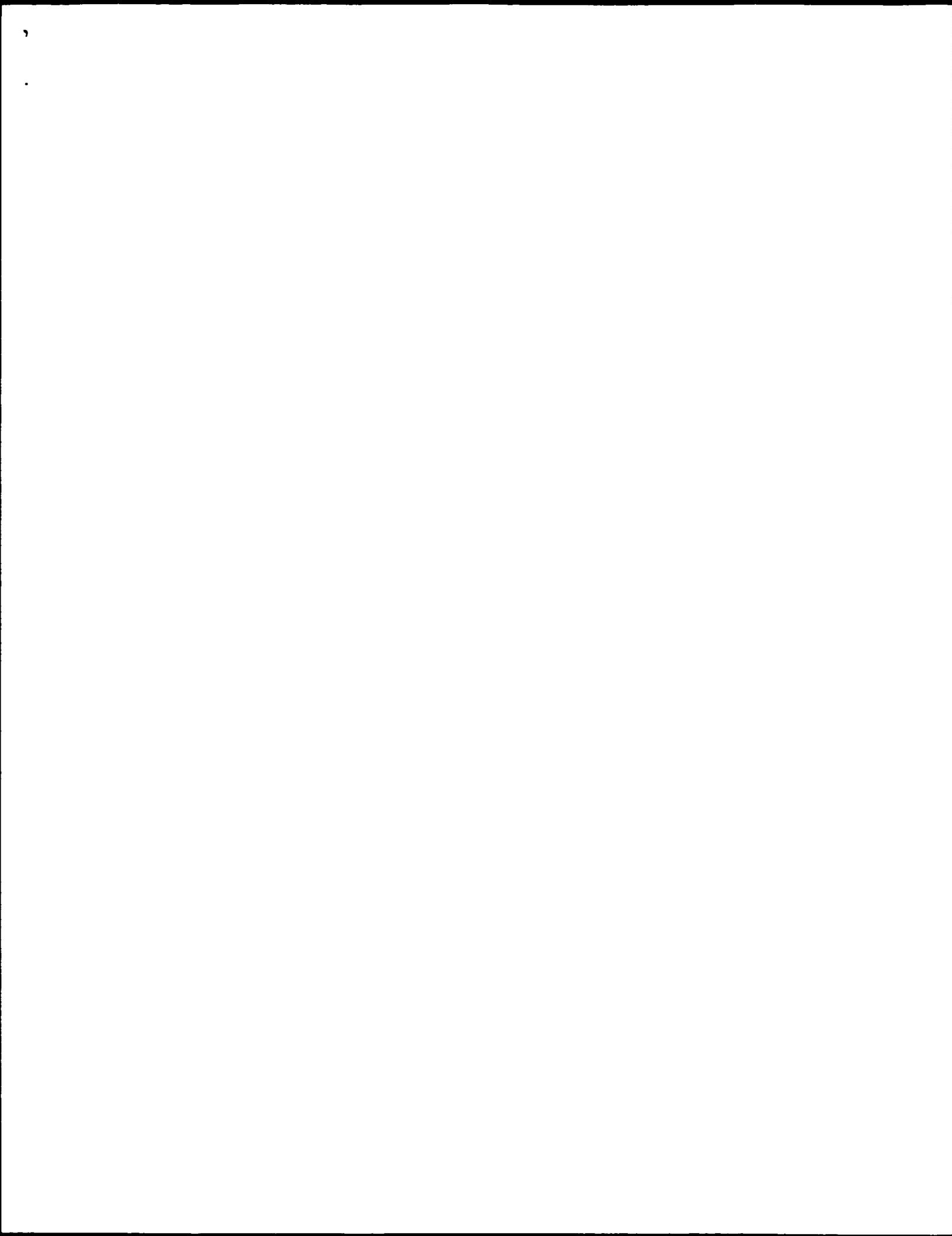
GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2002, 05:26:39 : Search time 22.2973 seconds
                                               (without alignments)
                                               727.075 Million cell updates/sec

Title: US-09-766-113-2
Perfect score: 66
Sequence: 1 tgagatctggatccattcga..........ccgtcctacgaattcagctg 66

Scoring table: IDENTITY_NUC
Database: Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Issued_Patents_NA:*
                 1:/cgn2_6/ptodata/1/ina/5A_COMB.seq:*
                 2:/cgn2_6/ptodata/1/ina/5B_COMB.seq:*
                 3:/cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                 3:/cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                 4:/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 5:/cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                 6:/cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB ID	Description
1	24.2	36.7	2290	1	US-08-584-226-1	Sequence 1, Appli
c 2	24.2	36.7	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c 3	22.4	33.9	3549	4	US-09-381-862-1	Sequence 1, Appli
c 4	22.2	33.6	2700	1	US-08-484-105-5	Sequence 5, Appli
c 5	22.2	33.6	2700	1	US-08-484-106-5	Sequence 5, Appli
c 6	22.2	33.6	443676	3	US-09-356-952-12	Sequence 12, Appl
c 7	22	33.3	1810	4	US-09-714-918-32	Sequence 32, Appl
c 8	22	33.3	1810	4	US-09-265-315-32	Sequence 32, Appl
c 9	22	33.3	1810	4	US-09-265-315-32	Sequence 32, Appl
c 10	22	33.3	1810	4	US-09-266-417-32	Sequence 32, Appl
c 11	21.8	33.0	2470	4	US-09-091-725-18	Sequence 18, Appl
c 12	21.8	33.0	2546	4	US-09-091-725-12	Sequence 12, Appl
c 13	21.8	33.0	3550	4	US-09-266-417-32	Sequence 22, Appl
c 14	21.6	32.7	730	1	US-07-826-945A-10	Sequence 10, Appl
c 15	21.6	32.7	1065	2	US-08-444-646-5	Sequence 5, Appl
c 16	21.6	32.7	1558	2	US-08-444-646-5	Sequence 5, Appl
c 17	21.2	32.1	4411529	4	US-09-103-840A-1	Sequence 4, Appli
c 18	21	31.8	1380	2	US-08-748-947A-1	Sequence 1, Appli
c 19	21	31.8	9432	1	US-08-277-231A-1	Sequence 1, Appli
c 20	21	31.8	9432	2	US-08-473-750-4	Sequence 4, Appli
c 21	21	31.8	9432	2	US-08-477-326-4	Sequence 4, Appli
c 22	20.8	31.5	2504	1	US-08-484-105-15	Sequence 15, Appl
c 23	20.8	31.5	2504	1	US-08-484-106-15	Sequence 15, Appl
c 24	20.8	31.5	2504	1	US-08-484-106-15	Sequence 15, Appl
c 25	20.8	31.5	3980	1	US-08-233-008A-1	Sequence 1, Appli
c 26	20.8	31.5	6519	1	US-08-233-008A-1	Sequence 1, Appli
c 27	20.6	31.2	9100	2	US-08-741-637B-27	Sequence 27, Appl
c 28	20.6	31.2	9100	3	US-08-526-840B-27	Sequence 27, Appl
c 29	20.4	30.9	230	2	US-08-532-896-28	Sequence 28, Appli
c 30	20.4	30.9	303	4	US-08-867-902F-2	Sequence 2, Appli
c 31	20.4	30.9	423	1	US-08-470-179-85	Sequence 85, Appl
c 32	20.4	30.9	472	2	US-08-465-380-33	Sequence 33, Appl
c 33	20.4	30.9	472	2	US-08-486-397-33	Sequence 33, Appl
c 34	20.4	30.9	472	2	US-08-486-399-33	Sequence 33, Appl
c 35	20.4	30.9	472	2	US-08-461-965-33	Sequence 33, Appl
c 36	20.4	30.9	472	2	US-08-634-641-33	Sequence 33, Appl
c 37	20.4	30.9	472	2	US-09-249-471-33	Sequence 33, Appl
c 38	20.4	30.9	472	2	US-09-249-451-33	Sequence 33, Appl
c 39	20.4	30.9	472	3	US-09-249-472-33	Sequence 33, Appl
c 40	20.4	30.9	472	3	US-08-809-455-33	Sequence 33, Appl
c 41	20.4	30.9	472	3	US-09-249-461-33	Sequence 33, Appl
c 42	20.4	30.9	472	3	US-09-249-448-33	Sequence 33, Appl
c 43	20.4	30.9	585	4	US-08-867-902F-3	Sequence 3, Appli
c 44	20.4	30.9	1407	3	US-08-688-988-7	Sequence 7, Appli
c 45	20.4	30.9	2156	4	US-08-965-762-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-08-584-226-1
Sequence 1, Application US/08584226
Patent No. 5798240

GENERAL INFORMATION:
APPLICANT: Martinis, Susan A.
APPLICANT: Sassanfar, Mandana
APPLICANT: Kim, Sunghoon
APPLICANT: Lee, Sang Ho
APPLICANT: Schimmel, Paul R.
TITLE OF INVENTION: RECOMBINANT MYCOBACTERIAL METHIONYL-tRNA
TITLE OF INVENTION: SYNTHETASE GENES, TESTER STRAINS AND ASSAYS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/584,226
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 536
APPLICATION NUMBER: US 08/305,766
FILING DATE: 13-SEP-1994
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI94-052
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2290 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1563
US-08-584-226-1

Query Match 96.7%; Score 24.2; DB 1; Length 2290;
Best Local Similarity 62.3%; Pred. No. 1.7;
Matches 48; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

```
Qy    6 TCTGATCCGTTCTGACTAAAATCGGTAAAAAAGTCGGTAGATTACCGTCCTACGAATTCAGCT 65
          | | | |  | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1428 TTTGGACTCGTCTCGGGCCCAGGGCCCAAAACCACGGTCGTTCGCCGCCCGTAGGTCGTTCGGGCT 1487

Qy   66 G 66
Db 1488 G 1488
```

RESULT 2
US-09-104-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328
GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
FILE REFERENCE: 24366-20007.00
CURRENT APPLICATION NUMBER: US/09/104,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g

Query Match 96.7%; Score 24.2; DB 4; length 4403765;
Best Local Similarity 62.3%; Pred. No. 9.4;
Matches 48; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

```
Qy       6 TCTGATCCGTTCTGACTAAAATCGGTAAAAAAGTCGGTAGATTACCGTCCTACGAATTCAGCT 65
             | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1125606 TTTGAACTCGTCTCGGTCCTAGGTCTCGGAAAAAATGGTAAAAAAACGAGCGGTCTGTTCGTTTCCGGCCCGTAGGTCGTTCGGGCT 1125547

Qy      66 G 66
Db 1125546 G 1125546
```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.40
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/381,862
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 1997-71077
FILING DATE: 25-MAR-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP98/01288
FILING DATE: 23-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Cawley, Jr., Thomas A.
REGISTRATION NUMBER: 40,944
REFERENCE/DOCKET NUMBER: 19036/46274
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3549 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Streptococcus pyogenes
STRAIN: Clinical Isolate SP-6-28
US-09-381-862-1

Query Match 33.9%; Score 22.4; DB 4; Length 3549;
Best Local Similarity 59.4%; Pred. No. 9.6;
Matches 38; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

```
Qy       1 TCGAGTCTGGATCCGTTCGGACAAAAACGGTAAAAAAGCGGTAGATTACCGTCCTACGAATT 60
             | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 3046 TCAGACCCAGGTTCGTTCGTTCTACAAGAACTTTCGACAAGGGTGAGTTCAGCGTTGTACCAGCTT 4105

Qy      61 CAGC 64
Db 3106 TACC 3109
```

RESULT 4
US-08-484-105-5/c
Sequence 5, Application US/08484105
Patent No. 5589341
GENERAL INFORMATION:
APPLICANT: STILLMAN, Bruce
APPLICANT: BELL, Stephen P
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: McNALLY, Francis J
APPLICANT: LAUKENSON, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

RESULT 4
US-09-481-862-1
Sequence 1, Application US/09381862
Patent No. 6245906
GENERAL INFORMATION:
APPLICANT: Ueyama, Hiroshi
APPLICANT: Abe, Kanako
APPLICANT: Keshi, Hitoyuki
APPLICANT: Matsushisa, Akio
TITLE OF INVENTION: PROBES FOR THE DIAGNOSIS OF INFECTIONS
TITLE OF INVENTION: CAUSED BY STREPTOCOCCUS PYOGENES
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive/6300 Sears Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606
COMPUTER READABLE FORM:

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,105
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2700 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-484-105-5

Qy   4 GAICTCGAICCGTTCGACAAAAACGGTAAAAAAAGCGTTAGATTACCTTCCTACGGAATTCA 62
        ||| | || |||||||| |||||||| ||| ||||| || || | ||||||| ||
Db 1586 GCTGAGAATGAATTCCACAAAAACGGATTCGAAAGATTTAGTTTGCCATCAACCGTATCCA 1528

Query Match   33.6%; Score 22.2; DB 1; Length 2700;
Best Local Similarity   61.0%; Pred. No. 11;
Matches 36; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

RESULT 5
US-08-484-106-5/c
Sequence 5, Application US/08484106
Patent No. 5614618
GENERAL INFORMATION:
APPLICANT: STILLMAN, Bruce
APPLICANT: BELL, Stephen P
APPLICANT: KOBAYASHI, Ryuji
APPLICANT: RINE, Jasper
APPLICANT: FOSS, Margit
APPLICANT: McNALLY, Francis J
APPLICANT: LAURENSON, Patricia
APPLICANT: HERSKOWITZ, Ira
APPLICANT: LI, Joachim J
APPLICANT: GAVIN, Kimberly
TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,106
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman Ph.D., Richard Aron
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-59032/DJB/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 494-8700
TELEFAX: (415) 494-8771
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2700 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-484-106-5

Qy   4 GATCTGGATCCGTTCGACAAAAACGGTAAAAAAGGGTAGAATTACCGTCCTACGGAATTCA 62
        | || ||| |||||||| |||||||| ||| ||||| || || | ||||||| ||
Db 1586 GCTGAGAATGAATTCCACAAAAACGGATTCGAAAGATTTAGTTTGCCATCAACCGTATCCA 1528

Query Match   33.6%; Score 22.2; DB 1; Length 2700;
Best Local Similarity   61.0%; Pred. No. 11;
Matches 36; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

RESULT 6
US-09-356-952-12/c
Sequence 12, Application US/09356952
Patent No. 6117663
GENERAL INFORMATION:
APPLICANT: Borlack-Sjodin, Ann
APPLICANT: Margarit, S. M.
APPLICANT: Bor-Soqi, Dafna
APPLICANT: Cole, Philip
APPLICANT: Kuriyan, John
TITLE OF INVENTION: A CRYSTAL OF A RAS-SOS COMPLEX AND METHODS OF USE
TITLE OF INVENTION: THEREOF
FILE REFERENCE: 600-1-228N
CURRENT APPLICATION NUMBER: US/09/356,952
CURRENT FILING DATE: 1999-07-19
EARLIER APPLICATION NUMBER: 60/093,631
EARLIER FILING DATE: 1998-07-21
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 43676
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
US-09-356-952-12

Qy   4 GATCTGGATCCGTTCGACAAAAACGGTAAAAAAGGGTAGAATTACCGTCCTACGGAATTCA   62
Db 30652 GCTGAGAATGAATTCCACAAAAACGGATCTGAAACGATTTAGTTTGCCATCAACCGTATCCA 30594

Query Match   33.6%; Score 22.2; DB 4; Length 43676;
Best Local Similarity   61.0%; Pred. No. 21;
Matches 36; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

RESULT 7
US-08-714-918-32/c
Sequence 32, Application US/08714918
Patent No. 6037123
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS ANTIBACTERIAL
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.

Left column

```
;  ZIP:  90071-2066
   COMPUTER READABLE FORM:
;    MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;    MEDIUM TYPE:  storage
;    COMPUTER:  IBM Compatible
;    OPERATING SYSTEM:  IBM P.C., DOS 5.0
;    SOFTWARE:  Word Perfect 5.1
   CURRENT APPLICATION DATA:
;    APPLICATION NUMBER:  US/08/714,918
;    FILING DATE:  September 13, 1996
;    CLASSIFICATION:  424
   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  60/009,102
;    FILING DATE:  December 22, 1995
;    APPLICATION NUMBER:  60/003,798
;    FILING DATE:  September 15, 1995
   ATTORNEY/AGENT INFORMATION:
;    NAME:  Warburg, Richard J.
;    REGISTRATION NUMBER:  32,327
;    REFERENCE/DOCKET NUMBER:  222/005
   TELECOMMUNICATION INFORMATION:
;    TELEPHONE:  (213) 489-1600
;    TELEFAX:  (213) 955-0440
;    TELEX:  67-3510
   INFORMATION FOR SEQ ID NO: 32:
   SEQUENCE CHARACTERISTICS:
;    LENGTH:  1810 base pairs
;    TYPE:  nucleic acid
;    STRANDEDNESS:  single
;    TOPOLOGY:  linear
   US-08-714-918-32
```

```
Query Match           33.4%;  Score 22;  DB 4;  Length 1810;
Best Local Similarity 61.8%;  Pred. No. 12;
Matches 34; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy     5 ATCTGGATCCGTTCGACAAAAACGGGTAAAAAAGCGGTAGATTACCGTCCTACGAAT   59
         ||||| ||||| || | | || | |||||| |||| ||| || ||||| || | |
Db  1669 ATCTGAATCCTTTTCACNAAAAGTAAAAGAAGGTAAAAGAAGGTAAAAGCTACCCTCGTATCATT 1615
```

```
RESULT 8
US-09-265-315-32/c
Sequence 32, Application US/09265315
Patent No. 6187541
GENERAL INFORMATION:
;    APPLICANT:  Benton, Bret
;    APPLICANT:  Lee, Ving J.
;    APPLICANT:  Malouin, Francois
;    APPLICANT:  Martin, Patrick K.
;    APPLICANT:  Schmid, Molly B.
;    APPLICANT:  Sun, Dongxu
   TITLE OF INVENTION:  METHODS OF SCREENING FOR COMPOUNDS
   TITLE OF INVENTION:  ACTIVE ON STAPHYLOCOCCUS AUREUS
   TITLE OF INVENTION:  TARGET GENES
;    NUMBER OF SEQUENCES:  111
   CORRESPONDENCE ADDRESS:
;    ADDRESSEE:  Lyon & Lyon
;    STREET:  633 West Fifth Street
;    STREET:  Suite 4700
;    CITY:  Los Angeles
;    STATE:  California
;    COUNTRY:  U.S.A.
;    ZIP:  90071-2066
   COMPUTER READABLE FORM:
;    MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;    MEDIUM TYPE:  storage
;    COMPUTER:  IBM Compatible
;    OPERATING SYSTEM:  IBM P.C., DOS 5.0
;    SOFTWARE:  Word Perfect 5.1
   CURRENT APPLICATION DATA:
;    APPLICATION NUMBER:  US/09/265,315
;    FILING DATE:  March 9, 1999
```

Right column

```
;  CLASSIFICATION:  435
   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  08/714,918
;    FILING DATE:  September 13, 1996
;    APPLICATION NUMBER:  60/009,102
;    FILING DATE:  December 22, 1995
;    APPLICATION NUMBER:  60/003,798
;    FILING DATE:  September 15, 1995
   ATTORNEY/AGENT INFORMATION:
;    NAME:  Warburg, Richard J.
;    REGISTRATION NUMBER:  32,327
;    REFERENCE/DOCKET NUMBER:  240/247
   TELECOMMUNICATION INFORMATION:
;    TELEPHONE:  (213) 489-1600
;    TELEFAX:  (213) 955-0440
;    TELEX:  67-3510
   INFORMATION FOR SEQ ID NO: 32:
   SEQUENCE CHARACTERISTICS:
;    LENGTH:  1810 base pairs
;    TYPE:  nucleic acid
;    STRANDEDNESS:  single
;    TOPOLOGY:  linear
   US-09-265-315-32
```

```
Query Match           33.4%;  Score 22;  DB 4;  Length 1810;
Best Local Similarity 61.8%;  Pred. No. 12;
Matches 34; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy     5 ATCTGGATCCGTTCGACAAAAACGGGTAAAAAAGCGGTAGATTACCGTCCTACGAAT   59
         ||||| ||||| || | | || | |||||| |||| ||| || ||||| || | |
Db  1669 ATCTGAATCCTTTTCACNAAAAGTAAAAGAAGGTAAAAGAAGGTAAAAGCTACCCTCGTATCATT 1615
```

```
RESULT 9
US-09-265-315-32/c
Sequence 32, Application US/09265315
Patent No. 6187541
GENERAL INFORMATION:
;    APPLICANT:  Benton, Bret
;    APPLICANT:  Lee, Ving J.
;    APPLICANT:  Malouin, Francois
;    APPLICANT:  Martin, Patrick K.
;    APPLICANT:  Schmid, Molly B.
;    APPLICANT:  Sun, Dongxu
   TITLE OF INVENTION:  METHODS OF SCREENING FOR COMPOUNDS
   TITLE OF INVENTION:  ACTIVE ON STAPHYLOCOCCUS AUREUS
   TITLE OF INVENTION:  TARGET GENES
;    NUMBER OF SEQUENCES:  111
   CORRESPONDENCE ADDRESS:
;    ADDRESSEE:  Lyon & Lyon
;    STREET:  633 West Fifth Street
;    STREET:  Suite 4700
;    CITY:  Los Angeles
;    STATE:  California
;    COUNTRY:  U.S.A.
;    ZIP:  90071-2066
   COMPUTER READABLE FORM:
;    MEDIUM TYPE:  3.5" diskette, 1.44 Mb
;    MEDIUM TYPE:  storage
;    COMPUTER:  IBM Compatible
;    OPERATING SYSTEM:  IBM P.C., DOS 5.0
;    SOFTWARE:  Word Perfect 5.1
   CURRENT APPLICATION DATA:
;    APPLICATION NUMBER:  US/09/265,315
;    FILING DATE:  March 9, 1999
   CLASSIFICATION:  435
   PRIOR APPLICATION DATA:
;    APPLICATION NUMBER:  08/714,918
;    FILING DATE:  September 13, 1996
;    APPLICATION NUMBER:  60/009,102
;    FILING DATE:  December 22, 1995
;    APPLICATION NUMBER:  60/003,798
;    FILING DATE:  September 15, 1995
```

RESULT 10
US-09-266-417-32/c
Sequence 32, Application US/09266417
Patent No. 6228588
GENERAL INFORMATION:
APPLICANT: Benton, Bret
APPLICANT: Lee, Ving J.
APPLICANT: Malouin, Francois
APPLICANT: Martin, Patrick K.
APPLICANT: Schmid, Molly B.
APPLICANT: Sun, Dongxu
TITLE OF INVENTION: METHODS OF SCREENING FOR COMPOUNDS
TITLE OF INVENTION: ACTIVE ON STAPHYLOCOCCUS AUREUS
TITLE OF INVENTION: TARGET GENES
NUMBER OF SEQUENCES: 111
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: Word Perfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/266,417
FILING DATE: March 9, 1999
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/714,918
FILING DATE: September 13, 1996
APPLICATION NUMBER: 60/009,102
FILING DATE: December 22, 1995
APPLICATION NUMBER: 60/003,798
FILING DATE: September 15, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 240/248
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1810 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-265-315-32

Qy　　　33.3%;　　Score 22;　DB 4;　Length 1810;
Best Local Similarity　61.8%;　Pred. No. 12;
Matches　34;　Conservative　0;　Mismatches　21;　Indels　0;　Gaps　0;

Query Match
US-09-265-315-32

Qy　5　ATCTGGATCCGTTCGACAAAACCGGTAAAAAACGGGTAGATTACCGTCGTACGGAAT 59
　　　|||||　|||||　||||||||||||| |||| | |||| || || ||| |||| |
Db　1669　ATCTGAATCCTTTCACNAAAAGTAAAAGAANGGTAAAAGCTACCCTCGTATCATT 1615

RESULT 11
US-09-091-725-18
Sequence 18, Application US/0909172 5
Patent No. 6329141
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Improved methods for transforming Phaffia
TITLE OF INVENTION: and recombinant DNA for use therein
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster llp
STREET: 2000 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: DC
COUNTRY: United States of America
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/091,725
FILING DATE: 23-DEC-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 95203620.0
FILING DATE: 22-DEC-1995
APPLICATION NUMBER: EP 96200943.7
FILING DATE: 11-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: E. Victor Donahue
REGISTRATION NUMBER: 35,492
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 2470 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Phaffia rhodozyma
FEATURE:
NAME/KEY: CDS
LOCATION: 177..2198
OTHER INFORMATION: /product= "pRcrty"
US-09-091-725-18

Qy　　　33.0%;　　Score 21.8;　DB 4;　Length 2470;
Best Local Similarity　61.4%;　Pred. No. 15;
Matches　35;　Conservative　0;　Mismatches　22;　Indels　0;　Gaps　0;

Query Match
US-09-091-725-18

Qy　1　TGAGTCTCGGATCCGTTCGACACAAAACCGGTAAAAAACGGTAAAAAGCCGGTAGATTACCGTCGTCCTACGA 57
　　　|||| | |||| |||　|||||||||　|| || | |||||| ||| ||| |||||| ||　||||| || |

Db 2168 TTTCGTTATGAGCGGATCGGAAGCGCACTAAGCACTCGGAAGAATACCGGACAGACAA 2224

RESULT 12
US 09/091-725-12
Sequence 12, Application US/09091725
Patent No. 6329141
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Improved methods for transforming Phaffia
; TITLE OF INVENTION: and recombinant DNA for use therein
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster llp
; STREET: 2000 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,725
; FILING DATE: 24-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95201620.0
; FILING DATE: 22-DEC-1995
; APPLICATION NUMBER: EP 96200943.7
; FILING DATE: 11-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: E. Victor Donahue
; REGISTRATION NUMBER: 45,492
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2546 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Phaffia rhodozyma
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 225..2246
; OTHER INFORMATION: /product= "PRGtTB"

US-09-091-725-12

Query Match 33.0%; Score 21.8; DB 4; Length 2546;
Host Local Similarity 61.4%; Pred. No. 15;
Matches 35; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy   1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGTAAAACCGTACATTACCGTCCTACGA 57
          ||||| ||||||||| |||||| |||||| |||||||| ||| |||||| |||| ||||||||
Db 2216 TGTCGTTATGAGCGGATCGGAAGCGCAGCGGAAGGCACTAAGCACAGCGGAAGAATACCGGACAGACAA 2272

RESULT 13
US-09-091-725-22
Sequence 22, Application US/09091725
Patent No. 6329141
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Improved methods for transforming Phaffia
; TITLE OF INVENTION: and recombinant DNA for use therein
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster llp
; STREET: 2000 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: United States of America
; ZIP: 20006-1888
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/091,725
; FILING DATE: 24-DEC-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95201620.0
; FILING DATE: 22-DEC-1995
; APPLICATION NUMBER: EP 96200943.7
; FILING DATE: 11-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: E. Victor Donahue
; REGISTRATION NUMBER: 45,492
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3550 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Phaffia rhodozyma
; STRAIN: CBS 6938
; FEATURE:
; NAME/KEY: exon
; LOCATION: 941..966
; FEATURE:
; NAME/KEY: intron
; LOCATION: 967..1077
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1078..1284
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1285..1364
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1365..1877
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1878..1959
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1960..2202
; FEATURE:
; NAME/KEY: intron
; LOCATION: 2203..2292
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2293..3325
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(941..966, 1078..1284, 1365..1877, 1960..2202, 2293..3325)
; OTHER INFORMATION: /product= "PRGtTB GH"

US-09-091-725-22

Query Match 33.0%; Score 21.8; DB 4; Length 3550;
Host Local Similarity 61.4%; Pred. No. 16;
Matches 35; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy   1 TGAGATCTGGATCCGTTCGACAAAACGGTAAAAAAGCGTAAAACCGTACATTACCGTCCTACGA 57

Db 3295 TGTGGTCATGAGGCGGATGGGAAGGGCAGTAAGCACACGGGAAGGGAGAAATACCGACAGACAA 3351

RESULT 14
US-07-826-945A-10
Sequence 10, Application US/07826945A
Patent No. 5270179
GENERAL INFORMATION:
APPLICANT: Chatterjee, Deb K.
TITLE OF INVENTION: Cloning and Expression of T5 DNA
TITLE OF INVENTION: Polymerase Reduced In 3'-to-5' Exonuclease Activity
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/826,945A
FILING DATE: 19920128
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Lewis, James L.
REGISTRATION NUMBER: 24,732
REFERENCE/DOCKET NUMBER: 0942.2270004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 466-0800
TELEFAX: (202) 833-8716
TELEX: 248636 SSK
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 730 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: both
US-07-826-945A-10

Query Match 32.7%; Score 21.6; DB 1; Length 730;
Best Local Similarity 63.5%; Pred. No. 14;
Matches 33; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

Qy 15 GTTCGACAAAACCGTAAAAAGCGGTAGAGATTACCGTCCTACGAATTCAGCTG 66
Db 530 GGTCGTCGAGGATTGCAAAAGCTGGTCGATTACCGTCCTATACTGACCTG 581

RESULT 15
US-08-444-646-5
Sequence 5, Application US/0844646
Patent No. 5837263
GENERAL INFORMATION:
APPLICANT: Haake, David A.
APPLICANT: Shang, Ellen S.
TITLE OF INVENTION: Leptospira MEMBRANE PROTEINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Robbins, Berliner & Carson
STREET: 201 N. Figueroa Street, 5th Floor
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90012-2628
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,646
FILING DATE: 19-MAY-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Wong, Wean Khing
REGISTRATION NUMBER: 33,561
REFERENCE/DOCKET NUMBER: 5656-107
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 977-1001
TELEFAX: (213) 977-1003
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1065 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-444-646-5

Query Match 32.7%; Score 21.6; DB 2; Length 1065;
Best Local Similarity 75.0%; Pred. No. 15;
Matches 27; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

Qy 23 AAACGGTAAAAAAGCGGTAGAATTACCGTCCTACGAA 58
Db 594 AGAAGTAAAAAAAGCGCGTTAGTTCCAGTCCTGCGAA 629

Search completed: November 5, 2002, 06:19:06
Job time : 1892.3 secs